

Wed Aug 15 13:35:08 2001

us-09-372-036-29.ra1

Page 8

Search completed: August 15, 2001, 13:10:02  
Job time: 39 sec

---

INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-34

Query Match 61.4%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OOTAPKA 8  
|||||  
DB 1 OOTAPKA 7

RESULT 14  
US-08-482-847-34  
Sequence 34, Application US/08482847  
Patent No. 5556757

GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
TITLE OF INVENTION: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.

COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993

ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 51916/104/INBI

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5390

TELEFAX: (202)672-5399

TELEFAX: 904136

INFORMATION FOR SEQ ID NO: 34:

SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

US-08-482-847-34

Query Match 61.4%; Score 35; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OOTAPKA 8  
|||||  
DB 1 OOTAPKA 7

RESULT 15  
US-08-456-670B-25  
Sequence 25, Application US/08456670B  
Patent No. 5932415

GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, Siegfried  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWILLER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STERAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US

ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992

ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA

REGISTRATION NUMBER: 33,302

REFERENCE/DOCKET NUMBER: MERCK 1694D1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333

TELEFAX: 703-243-6410

TELEFAX: 64191

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Listeria monocytogenes

STRAIN: EGD

US-08-456-670B-25

Query Match 61.4%; Score 35; DB 2; Length 9;  
Best Local Similarity 77.8%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 OOTAPKAPT 10  
|||||  
DB 1 OOTTKAPT 9



APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria innocua  
US-08-456-670B-38

Query Match 68.4%; Score 39; DB 2; Length 11;  
Best Local Similarity 63.6%; Pred. No. 0.55;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 QOQTAPKAPT 11  
:11111111:  
Db 1 EQOTTAKAPT 11

RESULT 11  
US-08-816-693A-51  
Sequence 51, Application US/08816693A  
Patent No. 5874241  
GENERAL INFORMATION:  
APPLICANT: Takahashi, Joseph S  
APPLICANT: Turek, Fred W  
APPLICANT: Pinto, Lawrence H  
TITLE OF INVENTION: Clock Gene and Gene Product  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
STREET: Two Prudential Plaza, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/816,693A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: 5874241thrup, Thomas E  
REGISTRATION NUMBER: 33,268  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-616-5400  
TELEFAX: 312-616-5460  
INFORMATION FOR SEQ ID NO: 51:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 747 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-816-693A-51

Query Match 64.9%; Score 37; DB 2; Length 747;

Best Local Similarity 70.0%; Pred. No. 90;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QOQTAPKAPT 10  
:11111111:  
Db 699 QOQPPQAPT 708

RESULT 12  
US-08-885-291-51  
Sequence 51, Application US/08885291A  
Patent No. 6057125  
GENERAL INFORMATION:  
APPLICANT: Takahashi, Joseph S.  
APPLICANT: Turek, Fred W.  
APPLICANT: Pinto, Lawrence H.  
TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
FILE REFERENCE: 0290-5  
CURRENT APPLICATION NUMBER: US/08/885,291A  
CURRENT FILING DATE: 1997-06-30  
EARLIER APPLICATION NUMBER: 08/816,693  
EARLIER FILING DATE: 1997-03-13  
NUMBER OF SEQ ID NOS: 55  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 51  
LENGTH: 747  
TYPE: PRT  
ORGANISM: Mus musculus  
US-08-885-291-51

Query Match 64.9%; Score 37; DB 3; Length 747;  
Best Local Similarity 70.0%; Pred. No. 90;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 QOQTAPKAPT 10  
:11111111:  
Db 699 QOQPPQAPT 708

RESULT 13  
US-08-127-499A-34  
Sequence 34, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136

TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-26

Query Match 100.0%; Score 57; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 0.027;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 OOOTAPKAPTE 11  
|||||  
Db 291 OOOTAPKAPTE 301

RESULT 9  
US-08-456-670B-20  
Sequence 20. Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P. C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 421911.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6533  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
SPRAIN: EGD  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1-7  
OTHER INFORMATION: /note= "SOME OR ALL Xaa AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 17-23  
OTHER INFORMATION: /note= "SOME OR ALL Xaa AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
US-08-456-670B-20

Query Match 82.5%; Score 47; DB 2; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.054;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 OOOTAPKAPT 10  
|||||  
Db 8 OOOTAPKAPT 16

RESULT 10  
US-08-456-670B-38  
Sequence 38. Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P. C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:

```

NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: MILLER, WHITE, ZELANO & BRANIGAN, P.C.
STREET: 2200 CLARENDON BLVD., SUITE 1400
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: US
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,670B
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/412,227
FILING DATE: 27-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/075,248
FILING DATE: 11-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4239567.4
FILING DATE: 25-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4219111.4
FILING DATE: 11-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: HAMLET-KING, DIANA
REGISTRATION NUMBER: 33,302
REFERENCE/DOCKET NUMBER: MERCK 1694D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
TELEX: 64191
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 478 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Listeria monocytogenes
STRAIN: EGD
US-08-456-670B-40

Query Match 100.0%; Score 57; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 1 000TAPKAPTE 11
|||||
Db 291 000TAPKAPTE 301

RESULT 7
US-08-127-499A-26
: Sequence 26, Application US/08127499A
: Patent No. 5510264
: GENERAL INFORMATION:
: APPLICANT: VAN ALSTYNE, Diane
: APPLICANT: SHARMA, Lawrence Rajendra
: TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
: TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES
: NUMBER OF SEQUENCES: 40
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington
: STATE: D.C.

```

```

1 COUNTRY: USA
2 ZIP: 20007-5109
3
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: Floppy disk
6 COMPUTER: IBM PC compatible
7 OPERATING SYSTEM: PC-DOS/MS-DOS
8 SOFTWARE: Patentin Release #1.0, Version #1.30
9 CURRENT APPLICATION DATA:
10 APPLICATION NUMBER: US/08/127,499A
11 FILING DATE: 28-SEP-1993
12 ATTORNEY/AGENT INFORMATION:
13 NAME: BENT, Stephen A.
14 REGISTRATION NUMBER: 29,768
15 REFERENCE/DOCKET NUMBER: 51916/102/INBI
16 TELECOMMUNICATION INFORMATION:
17 TELEPHONE: (202)672-5300
18 TELEFAX: (202)672-5399
19
20 TELEFAX: (202)672-5399
21
22 INFORMATION FOR SEQ ID NO: 26:
23 SEQUENCE CHARACTERISTICS:
24 LENGTH: 484 amino acids
25 TYPE: amino acid
26 STRADEDNESS:
27 TOPOLOGY: unknown
28
29 US-08-127-499A-26
30
31 Query Match 100.0%; Score 57; DB 1; Length 484;
32 Best Local Similarity 100.0%; Pred. No. 0.027;
33 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
34
35 QY 1 000TAPKAPTE 11
36 |||||||||
37 Db 291 000TAPKAPTE 301
38
39 RESULT 8
40 US-08-482-847-26
41 Sequence 26, Application US/08482847
42 Patent No. 5556757
43 GENERAL INFORMATION:
44 APPLICANT: VAN ALSTYNE, Diane
45 TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR
46 BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR
47 TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR
48 NUMBER OF SEQUENCES: 40
49 CORRESPONDENCE ADDRESS:
50 ADDRESSEE: Foley & Lardner
51 STREET: 3000 K Street, N.W., Suite 500
52 CITY: Washington
53 STATE: D.C.
54 COUNTRY: USA
55 ZIP: 20007-5109
56 COMPUTER READABLE FORM:
57 MEDIUM TYPE: Floppy disk
58 COMPUTER: IBM PC compatible
59 OPERATING SYSTEM: PC-DOS/MS-DOS
60 SOFTWARE: Patentin Release #1.0, Version #1.30
61 CURRENT APPLICATION DATA:
62 APPLICATION NUMBER: US/08/482,847
63 FILING DATE: 07-JUN-1995
64 CLASSIFICATION: 514
65 PRIOR APPLICATION DATA:
66 APPLICATION NUMBER: US 08/127,499
67 FILING DATE: 28-SEP-1993
68 ATTORNEY/AGENT INFORMATION:
69 NAME: BENT, Stephen A.
70 REGISTRATION NUMBER: 29,768
71 REFERENCE/DOCKET NUMBER: 51916/104/INBI
72 TELECOMMUNICATION INFORMATION:
73 TELEPHONE: (202)672-5300
74 TELEFAX: (202)672-5399
75

```

RESULT 4  
US-08-482-847-33  
; Sequence 33, Application US/08482847  
; Patent No. 5556757  
; GENERAL INFORMATION:  
; APPLICANT: VAN ALSTYNE, Diane  
; APPLICANT: SHARMA, Lawrence Rajendra  
; TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
; TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
; TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,847  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/127,499  
; FILING DATE: 28-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 51916/104/INBI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acid  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; US-08-482-847-33

Query Match 100.0%; Score 57; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 000TAPKAPTE 11  
|||||  
Db 7 000TAPKAPTE 17

RESULT 5  
US-08-456-670B-39  
; Sequence 39, Application US/08456670B  
; Patent No. 5932415  
; GENERAL INFORMATION:  
; APPLICANT: SCHUBERT, PETER  
; APPLICANT: NEUMANN, SIEGFRIED  
; APPLICANT: PAMELZIK, MARTINA  
; APPLICANT: LINKEWELER, WINFRIED  
; APPLICANT: BURGER, CHRISTA  
; APPLICANT: HOFMANN, GOTTFRIED  
; APPLICANT: HUBERT, ANDREAS  
; APPLICANT: GOEBEL, WERNER  
; APPLICANT: KOHLER, STEFAN  
; TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
; TITLE OF INVENTION: LISTERIAS  
; NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
; STREET: 2200 CLARENDON BLVD., SUITE 1400  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: US  
; ZIP: 22201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,670B  
; FILING DATE: 01-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/412,227  
; FILING DATE: 27-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/075,248  
; FILING DATE: 11-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 4239567.4  
; FILING DATE: 25-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 421911.4  
; FILING DATE: 11-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HAMELT-KING, DIANA  
; REGISTRATION NUMBER: 33,302  
; REFERENCE/DOCKET NUMBER: MERCK 1694D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-243-6333  
; TELEFAX: 703-243-6410  
; TELEX: 64191  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 232 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Listeria monocytogenes  
; STRAIN: EGD  
; US-08-456-670B-39

Query Match 100.0%; Score 57; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 000TAPKAPTE 11  
|||||  
Db 222 000TAPKAPTE 232

RESULT 6  
US-08-456-670B-40  
; Sequence 40, Application US/08456670B  
; Patent No. 5932415  
; GENERAL INFORMATION:  
; APPLICANT: SCHUBERT, PETER  
; APPLICANT: NEUMANN, SIEGFRIED  
; APPLICANT: PAMELZIK, MARTINA  
; APPLICANT: LINKEWELER, WINFRIED  
; APPLICANT: BURGER, CHRISTA  
; APPLICANT: HOFMANN, GOTTFRIED  
; APPLICANT: HUBERT, ANDREAS  
; APPLICANT: GOEBEL, WERNER  
; APPLICANT: KOHLER, STEFAN  
; TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
; TITLE OF INVENTION: LISTERIAS

TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-29

Query Match 100.0%; Score 57; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00055;  
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOOTAPKAPTE 11  
Db 1 QOOTAPKAPTE 11

RESULT 2  
US-08-456-670B-42  
Sequence 42, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, STEGRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWILLER, WINFRIED  
APPLICANT: BUNGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUEBEL, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 421911.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694DI

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-456-670B-42

Query Match 100.0%; Score 57; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0006;  
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOOTAPKAPTE 11  
Db 2 QOOTAPKAPTE 12

RESULT 3  
US-08-127-499A-33  
Sequence 33, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-33

Query Match 100.0%; Score 57; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOOTAPKAPTE 11  
Db 7 QOOTAPKAPTE 17

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 13:10:02 ; Search time 30.68 Seconds  
(without alignments)  
7.382 Million cell updates/sec

Title: US-09-372-036-29

Perfect score: 57

Sequence: 1 000TAPKAPTE 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues  
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/PCtUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	11	2	US-08-456-670B-29
2	57	100.0	12	2	US-08-456-670B-42
3	57	100.0	21	1	US-08-127-459A-33
4	57	100.0	21	1	US-08-482-847-33
5	57	100.0	23	2	US-08-456-670B-39
6	57	100.0	478	2	US-08-456-670B-40
7	57	100.0	484	1	US-08-127-459A-26
8	57	100.0	484	1	US-08-482-847-26
9	47	82.5	23	2	US-08-456-670B-20
10	39	68.4	11	2	US-08-456-670B-38
11	37	64.9	747	2	US-08-816-693A-51
12	37	64.9	747	2	US-08-885-291-51
13	35	61.4	7	1	US-08-127-459A-34
14	35	61.4	7	1	US-08-482-847-34
15	35	61.4	9	2	US-08-456-670B-25
16	35	61.4	443	1	US-08-342-411A-4
17	34	59.6	405	5	PCT-US93-1140A-2
18	34	59.6	3457	2	US-08-416-603-4
19	33	57.9	21	2	US-08-934-915-72
20	32	56.1	73	3	US-09-041-889-12
21	32	56.1	73	3	US-08-837-058-12
22	32	56.1	220	3	US-09-041-889-2
23	32	56.1	220	3	US-08-837-058-2
24	32	56.1	298	4	US-08-961-083-24
25	32	56.1	459	4	US-09-118-319-6
26	32	56.1	459	4	US-09-286-691-2
27	32	56.1	459	4	US-09-687-147-2

28	32	56.1	605	2	US-08-687-956A-1	Sequence 1, Appl1
29	32	56.1	696	6	5262177-4	Patent No. 5262177
30	32	56.1	698	2	US-08-175-158A-2	Sequence 2, Appl1
31	32	56.1	788	2	US-08-918-914-4	Sequence 2, Appl1
32	32	56.1	903	2	US-08-853-310-2	Sequence 2, Appl1
33	32	56.1	1074	2	US-08-470-058-2	Sequence 2, Appl1
34	32	56.1	1074	2	US-09-037-188-2	Sequence 2, Appl1
35	32	56.1	1162	2	US-08-728-333A-2	Sequence 2, Appl1
36	32	56.1	1410	2	US-08-470-058-4	Sequence 4, Appl1
37	32	56.1	1410	2	US-09-037-188-4	Sequence 4, Appl1
38	32	56.1	1572	2	US-08-290-731C-5	Sequence 5, Appl1
39	32	56.1	1596	3	US-09-356-952-3	Sequence 3, Appl1
40	32	56.1	3118	2	US-08-457-273B-8	Sequence 8, Appl1
41	31	54.4	196	2	US-08-481-814A-9	Sequence 9, Appl1
42	31	54.4	339	1	US-08-248-629A-3	Sequence 3, Appl1
43	31	54.4	339	1	US-08-451-932-3	Sequence 3, Appl1
44	31	54.4	339	1	US-08-452-260-3	Sequence 3, Appl1
45	31	54.4	339	2	US-08-612-788-3	Sequence 3, Appl1

## ALIGNMENTS

RESULT 1  
US-08-456-670B-29  
: Sequence 29, Application US/08456670B  
: Patent No. 5932415  
: GENERAL INFORMATION:  
: APPLICANT: SCHUBERT, PETER  
: APPLICANT: NEUMANN, Siegfried  
: APPLICANT: PAMELZIK, MARTINA  
: APPLICANT: LINKEWELER, Winfried  
: APPLICANT: BURGER, Christa  
: APPLICANT: HOFMANN, Gottfried  
: APPLICANT: BOBERT, Andreas  
: APPLICANT: GOEBEL, Werner  
: APPLICANT: KOHLER, Stefan  
: TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
: TITLE OF INVENTION: LISTERIAS  
: NUMBER OF SEQUENCES: 43  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: MILLEN, WHITE, ZELANO & ARANIGAN, P.C.  
: STREET: 2200 CLARENDON BLVD., SUITE 1400  
: CITY: ARLINGTON  
: STATE: VIRGINIA  
: COUNTRY: US  
: ZIP: 22201  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/456,670B  
: FILING DATE: 01-JUN-1995  
: CLASSIFICATION: 435  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/412,227  
: FILING DATE: 27-MAR-1995  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 08/075,248  
: FILING DATE: 11-JUN-1993  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: DE 4239567.4  
: FILING DATE: 25-NOV-1992  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: DE 4219111.4  
: FILING DATE: 11-JUN-1992  
: ATTORNEY/AGENT INFORMATION:  
: NAME: HAMLET-KING, DIANA  
: REGISTRATION NUMBER: 33,302  
: REFERENCE/DOCKET NUMBER: MERCK 169401  
: TELECOMMUNICATION INFORMATION:

***This Page Blank (usp10)***

```

DE GLUTAMINE SYNTHETASE (EC 6.3.1.2) (GLUTAMATE--AMMONIA LIGASE).
GN GLN OR SLR1756.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_Taxid-1148;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE-95244839; PubMed-7727755;
RA Reyes J.C., Florencio F.J.;
RT "Electron transport controls the transcription of the glutamine
synthetase gene (glnA) from the cyanobacterium Synechocystis sp. PCC
6803.";
RT Plant Mol. Biol. 27:789-799(1995).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE-97061201; PubMed-8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
Miyajima N., Hirosewa M., Suglura M., Sasamoto S., Kimura T.,
Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synechocystis sp. strain PCC6803. II. Sequence determination of the
entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE +
ORTHOPHOSPHATE.
CC -1- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO
HEXAGON (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X69199; CAA49139.1; -
DR EMBL: D90903; BAA17055.1; -
DR HSSP: P06201; 2LGS.
DR InterPro: IPR001691; -
DR Pfam: PF00120; gln-synt; 1.
DR PROSITE: PS00180; GLN_1; 1.
DR PROSITE: PS00181; GLN_ATP; 1.
KW Ligase.
FT CONFLICT 115 115 V -> A (IN REF. 2).
FT CONFLICT 145 147 QTE -> PNG (IN REF. 1).
FT CONFLICT 180 180 O -> E (IN REF. 1).
FT CONFLICT 204 206 GLC -> ARG (IN REF. 2).
FT CONFLICT 230 230 K -> KFDK (IN REF. 1).
FT CONFLICT 239 239 M -> I (IN REF. 1).
FT CONFLICT 358 358 K -> N (IN REF. 1).
FT CONFLICT 378 379 ML -> IV (IN REF. 1).
FT CONFLICT 430 430 E -> Q (IN REF. 1).
SO SEQUENCE 473 AA; 53026 MW; 75F8E28B5EA9515 CRC64;

```

```

Query Match 50.0%; Score 5; DB 1; Length 473;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 PVAPT 5
   |||||
DB 184 PVAPT 188

```



Query Match 50.0%; Score 5; DB 1; Length 453;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAAPT 5  
 DB 374 PVAAPT 378

RESULT 48  
 ID REBA\_VIBCH STANDARD; PRT; 465 AA.  
 AC 007024; 09KVA6;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PUTATIVE MANNOSE-1-PHOSPHATE GUANYLYLTRANSFERASE (GDP) (EC 2.7.7.22)  
 GN REBA OR VC0241.  
 OS *Vibrio cholerae*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.  
 OX NCBI\_Taxid=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-EL TOR 017 / SEROTYPE OGAWA;  
 RX MEDLINE-92212870; PubMed-1372980;  
 RA Strother U.H., Karageorgos L.E., Morona R., Manning P.A.;  
 RT "Serotype conversion in *Vibrio cholerae* O1."  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2565-2570(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-EL TOR N16961 / SEROTYPE O1;  
 RX MEDLINE-20406833; PubMed-10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tetelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Drygol I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*."  
 RL Nature 406:477-483(2000).  
 CC -1- CATALYTIC ACTIVITY: GDP + D-MANNOSE 1-PHOSPHATE = ORTHOPHOSPHATE + GDP-MANNOSE.  
 CC -1- PATHWAY: GDP-MANNOSE BIOSYNTHESIS WITHIN THE O ANTIGEN BIOSYNTHESIS PATHWAY OF LIPOPOLYSACCHARIDE BIOSYNTHESIS.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF MANNOSE-6-PHOSPHATE ISOMERASES.  
 CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 14.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X59554; CAA42134.1; ALT\_FRAME.  
 DR EMBL: AE004113; AAF93417.1; .  
 DR PIR: S28468; S28468.  
 DR TIGR: VC0241; .  
 DR InterPro: IPR001538; .  
 DR InterPro: IPR001825; .  
 DR Pfam: PF01050; MannoseP\_isomert; 1.  
 DR Pfam: PF00483; NTP\_transferase; 1.  
 KW Lipopolysaccharide biosynthesis; Transferase; Kinase.  
 SEQUENCE 465 AA; 51916 MW; 7274A5A876F49268 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 465;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 QEVKK 10  
 DB 337 QEVKK 341

RESULT 49  
 ID GLNA\_FREDI STANDARD; PRT; 470 AA.  
 AC P33035;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE GLUTAMINE SYNTHETASE (EC 6.3.1.2) (GLUTAMATE--AMMONIA LIGASE).  
 GN GLNA.  
 OS *Fremyella diplosiphon* (Calothrix PCC 7601).  
 OC Bacteria; Cyanobacteria; Nostocales; Rivulariaceae; *Fremyella*.  
 OX NCBI\_Taxid=1197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-93129187; PubMed-1362348;  
 RA Elmorjani K., Liotenberg S., Hounard J., de Marsac N.T.;  
 RT "Molecular characterization of the gene encoding glutamine synthetase in the cyanobacterium *Calothrix* sp. PCC 7601."  
 RL Biochem. Biophys. Res. Commun. 189:1296-1302(1992).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE + ORTHOPHOSPHATE.  
 CC -1- SUBUNIT: OLIGOMER OF 12 SUBUNITS ARRANGED IN THE FORM OF TWO HEXAGON.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L05609; AAA23288.1; .  
 DR PIR: JCI403; JCI403.  
 DR HSSP: P06201; ZLGS.  
 DR InterPro: IPR001637; .  
 DR InterPro: IPR001691; .  
 DR Pfam: PF00120; gln-synt; 1.  
 DR PROSITE: PS00180; GLNA\_1; 1.  
 DR PROSITE: PS00181; GLNA\_ATP; 1.  
 DR PROSITE: PS00182; GLNA\_ADENYLYATION; 1.  
 KW Ligase.  
 FT INIT\_MET 0 BY SIMILARITY.  
 SEQUENCE 470 AA; 52919 MW; ADD7B49A7789E832 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 470;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAAPT 5  
 DB 181 PVAAPT 185

RESULT 50  
 ID GLNA\_SYNY3 STANDARD; PRT; 473 AA.  
 AC P77961; 059981;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

CC -1- SIMILARITY: TO YEAST YBR287W.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z88056; CAB10852.1; -.  
CC Hypothetical protein; Transmembrane.  
CC TRANSMEM 18 38 POTENTIAL.  
CC TRANSMEM 81 101 POTENTIAL.  
CC TRANSMEM 269 289 POTENTIAL.  
CC TRANSMEM 317 337 POTENTIAL.  
CC TRANSMEM 354 374 POTENTIAL.  
CC TRANSMEM 390 410 POTENTIAL.  
CC TRANSMEM 428 448 POTENTIAL.  
CC SEQUENCE 452 AA; 49575 MW; 71B77EA5725C69A8 CRC64;  
CC -----  
Query Match 50.0%; Score 5; DB 1; Length 452;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 4 PROEV 8  
DB 342 PROEV 346  
CC -----  
RESULT 46  
GATA\_HELPY  
ID GATA\_HELPY STANDARD: PRT: 453 AA.  
AC 092L13;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE SUBUNIT A (EC 6.3.5.-) (GLU-ADT  
DE SUBUNIT A).  
GN GATA OR JHP0769.  
OS Helicobacter pylori J99 (Campylobacter pylori J99).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=85963;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99120557; PubMed=9923682;  
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,  
RA Smith D.R., Noonan B., Guild B.C., deJonghe B.L., Carmel G.,  
RA Tummiano P.J., Caruso A., Ulla-Nickelsen M., Mills D.M., Ives C.,  
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
RA Trust T.J.;  
RT "Genomic sequence comparison of two unrelated isolates of the human  
RT gastric pathogen Helicobacter pylori";  
RL Nature 397:176-180(1999).  
CC -1- FUNCTION: FURNISHES A MEANS FOR FORMATION OF CORRECTLY CHARGED  
CC GLN-TRNA(GLN) THROUGH THE TRANSMUTATION OF MISACTYLATED GLU-  
CC TRNA(GLN) IN ORGANISMS WHICH LACK GLUTAMINYL-TRNA SYNTHETASE. THE  
CC REACTION TAKES PLACE IN THE PRESENCE OF GLUTAMINE AND ATP THROUGH  
CC AN ACTIVATED GAMMA-PHOSPHO-GLU-TRNA(GLN) (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMYL-TRNA(GLN) + L-GLUTAMINE - ADP  
CC + PHOSPHATE + L-GLUTAMINYL-TRNA(GLN) + L-GLUTAMATE.  
CC -1- SUBUNIT: HETEROTRIMER OF A, B AND C SUBUNITS (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC -----  
CC EMBL: AE001507; AAD06348.1; -.  
CC InterPro: IPR000120; -.  
CC Pfam: PF01425; Amidase; 1.  
CC PROSITE: PS00571; AMIDASES; 1.  
CC Protein biosynthesis; Ligase.  
CC SEQUENCE 453 AA; 49749 MW; FD97731532E7397 CRC64;  
CC -----  
Query Match 50.0%; Score 5; DB 1; Length 453;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 PVAPT 5  
DB 374 PVAPT 378  
CC -----  
RESULT 47  
GATA\_HELPY  
ID GATA\_HELPY STANDARD: PRT: 453 AA.  
AC P56114;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE SUBUNIT A (EC 6.3.5.-) (GLU-ADT  
DE SUBUNIT A).  
GN GATA OR HP0830.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=26695 / ATCC 700392;  
RX MEDLINE=97394467; PubMed=9252185;  
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,  
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,  
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
RA Venter J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
RT pylori";  
RL Nature 388:539-547(1997).  
CC -1- FUNCTION: FURNISHES A MEANS FOR FORMATION OF CORRECTLY CHARGED  
CC GLN-TRNA(GLN) THROUGH THE TRANSMUTATION OF MISACTYLATED GLU-  
CC TRNA(GLN) IN ORGANISMS WHICH LACK GLUTAMINYL-TRNA SYNTHETASE. THE  
CC REACTION TAKES PLACE IN THE PRESENCE OF GLUTAMINE AND ATP THROUGH  
CC AN ACTIVATED GAMMA-PHOSPHO-GLU-TRNA(GLN) (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMYL-TRNA(GLN) + L-GLUTAMINE - ADP  
CC + PHOSPHATE + L-GLUTAMINYL-TRNA(GLN) + L-GLUTAMATE.  
CC -1- SUBUNIT: HETEROTRIMER OF A, B AND C SUBUNITS (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AE000594; AAD07880.1; -.  
CC TIGR: HP0830; -.  
CC InterPro: IPR000120; -.  
CC Pfam: PF01425; Amidase; 1.  
CC PROSITE: PS00571; AMIDASES; 1.  
CC Protein biosynthesis; Ligase.  
CC SEQUENCE 453 AA; 49652 MW; A35814B32F1AE13A CRC64;  
CC -----

```

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Ulla-Nikelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- CATALYTIC ACTIVITY: GLUCOSAMINE 1-PHOSPHATE = GLUCOSAMINE 6-
CC PHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOHEXOSE MUTASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001446; AAD05654.1; -
DR InterPro: IPR001485; -
DR Pfam: PF00408; PGM_PMM: 1.
DR PRINTS: PR00509; PGM_PMM: 1.
DR PROSITE: PS00710; PGM_PMM: 1.
KM Isomerase; Phosphorylation.
FT ACT_SITE 99
FT SEQUENCE 445 AA; 49174 MW; 586DD2E161295667 CRC64;
SQ
Query Match 50.0%; Score 5; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 6 OEYKK 10
DB 228 OEYKK 232
RESULT 44
GLMM_HELPY
ID GLMM_HELPY STANDARD; PRT: 445 AA.
AC P25177;
DT 01-MAY-1992 (rel. 22, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 01-OCT-2000 (rel. 40, Last annotation update)
DE PHOSPHOGLUCOSAMINE MUTASE (EC 5.4.2.-).
GN GLMM OR UREC OR HP0075.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=85P;
RX MEDLINE=91161505; PubMed=2001995;
RA Labigne A., Cussac V., Courcoux P.;
RT "Shuttle cloning and nucleotide sequences of Helicobacter pylori
RT genes responsible for urease activity."
RL J. Bacteriol. 173:1920-1931(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,

```

```

RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalaf H.G., Gloeck A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- CATALYTIC ACTIVITY: GLUCOSAMINE 1-PHOSPHATE = GLUCOSAMINE 6-
CC PHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOHEXOSE MUTASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60398; AAA25018.1; -
DR EMBL: AE000529; AAD07146.1; -
DR PIR: C38537; C38537.
DR TIGR: HP0075; -
DR InterPro: IPR001485; -
DR Pfam: PF00408; PGM_PMM: 1.
DR PROSITE: PS00710; PGM_PMM: 1.
KM Isomerase; Phosphorylation.
FT ACT_SITE 99
FT SEQUENCE 445 AA; 49086 MW; C0A52D904FFDAF20 CRC64;
SQ
Query Match 50.0%; Score 5; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 6 OEYKK 10
DB 228 OEYKK 232
RESULT 45
IDQ4_SCHPO
ID YDQ4_SCHPO STANDARD; PRT: 452 AA.
AC O14197;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 15-JUL-1998 (rel. 36, Last annotation update)
DE HYPOTHETICAL 49.6 KDA PROTEIN C5D6.04 IN CHROMOSOME I.
GN SPAC5D6.04.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX Skellern J., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).

```

```

CC -----
DR EMBL: AB013466; BAA33429.1; -
DR EMBL: AB013467; BAA33430.1; -
DR EMBL: AB013469; BAA33431.1; -
DR EMBL: AB013469; BAA33432.1; -
DR EMBL: AF079971; AAC77924.1; -
DR EMBL: U83896; AAB41444.1; -
DR MGI: 1334255; Pscd2.
DR HSSP: 099418; 1PBV.
DR InterPro: IPR000904; -
DR InterPro: IPR001849; -
DR Pfam: PF00169; PH; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
KW Guanine-nucleotide releasing factor: Colled coll.
KM Alternative splicing.
FT DOMAIN 10 67 COILED COIL (POTENTIAL).
FT ID 72 201 SECT.
FT DOMAIN 259 376 PH.
FT VARSPLIC 1 16 MISSING (IN ISOFORM 3).
FT VARSPLIC 273 273 MISSING (IN ISOFORM 2).
SO SEQUENCE 400 AA; 46571 MW; 88F14DA22A331A89 CRC64;

```

Query Match 50.0%; Score 5; DB 1; Length 400;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APPOE 7  
 Db 358 APPOE 362

```

RESULT 41
YLN2_CAEEL
ID YLN2_CAEEL STANDARD; PRT; 415 AA.
AC 018964;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHEICAL 46.2 KDA TRP-ASP REPEATS CONTAINING PROTEIN D2013.2 IN
DE CHROMOSOME II.
GN D2013.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN 11
RP SEQUENCE FROM N.A.
RA Mortimore B.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z47808; CAA87770.1; -
DR WormRep: D2013.2; CE00928.
DR InterPro: IPR000306; -
DR InterPro: IPR001680; -
DR Pfam: PF01363; FYVE; 1.
DR Pfam: PF00400; WD40; 5.
DR PRINTS: PR00320; GPROTEINRPT.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS50082; WD_REPEATS_2; 2.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat.
FT REPEAT 71 103 WD 1.

```

```

FT REPEAT 119 148 WD 2.
FT REPEAT 202 232 WD 3.
FT REPEAT 373 403 WD 4.
SO SEQUENCE 415 AA; 46239 MW; 8927612C040F5E43 CRC64;

```

Query Match 50.0%; Score 5; DB 1; Length 415;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
 Db 75 PVAPT 79

```

RESULT 42
SCK_HUMAN
ID SCK_HUMAN STANDARD; PRT; 428 AA.
AC P98077;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE PROTEIN SCK (FRAGMENT).
DE SCK.
GN SCK.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=95090462; PubMed=7527937;
RA Kavanaugh W.M., Williams L.T.;
RT "An alternative to SH2 domains for binding tyrosine-phosphorylated
RT proteins."
RL Science 266:1862-1865(1994).
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN LIVER, ALSO PRESENT IN BRAIN.
CC -1- SIMILARITY: CONTAINS 1 PID DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: TO SHC TRANSFORMING PROTEINS.
DE HSSP: P29353; IMIL.
DR InterPro: IPR000050; -
DR InterPro: IPR000980; -
DR Pfam: PF00640; PID; 1.
DR Pfam: PF00017; SH2; 1.
DR PROSITE: PS01179; PID; 1.
DR PROSITE: PS50001; SH2; 1.
KW SH2 domain.
FT NON_TER 1 1
FT DOMAIN 34 216 PID.
FT DOMAIN 374 >428 SH2.
FT NON_TER 428 428
SO SEQUENCE 428 AA; 46522 MW; D19A3FDECCB1BC6 CRC64;

```

Query Match 50.0%; Score 5; DB 1; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
 Db 361 PVAPT 365

```

RESULT 43
GLM_HELPJ
ID GLM_HELPJ STANDARD; PRT; 445 AA.
AC Q92M22;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PHOSPHOLUCOSAMINE MOTASE (EC 5.4.2.-).
GN GLM OR UREC OR JHP0070.
OS Helicobacter pylori J99 (Campylobacter pylori J99).

```

Db 180 PVAPT 184

RESULT 39

CYH2\_HUMAN STANDARD: PRT; 400 AA.

AC Q99418; Q92958; 15-JUL-1998 (Rel. 36, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE CYTOSKESTIN 2 (ARF NUCLEOTIDE-BINDING SITE OPENER) (ARNO PROTEIN) (ARF EXCHANGE FACTOR).

GN PSCD2 OR ARNO.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Brain;

RX MEDLINE=97100951; PubMed=8945478;

RA Charatin P., Patis S., Antony B., Robineau S., Bernaud-Dufour S., Jackson C.L., Chadre M.;

RT "A human exchange factor for ARF contains Sec7 and pleckstrin-homology domains.";

RL Nature 384:481-484(1996).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.

RC TISSUE=Brain;

RX MEDLINE=98079021; PubMed=9417041;

RA Frank S.F., Uppender S.K., Hansen S.H., Casanova J.F.;

RT "ARNO is a guanine nucleotide exchange factor for ADP-ribosylation factor 6.";

RL J. Biol. Chem. 273:23-27(1998).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 51-252.

RX MEDLINE=98135767; PubMed=9476900;

RA Mossesova E., Gulbis J.M., Goldberg J.;

RT "Structure of the guanine nucleotide exchange factor Sec7 domain of human ARNO and analysis of the interaction with ARF GTPase.";

RL Cell 92:415-423(1998).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 51-252.

RX MEDLINE=98169075; PubMed=9510256;

RA Cherfils J., Menetrey J., Mathieu M., Le Bras G., Robineau S., Bernaud-Dufour S., Antony B., Chardin P.;

RT "Structure of the Sec7 domain of the Arf exchange factor ARNO.";

RL Nature 392:101-105(1998).

CC -1- FUNCTION: PROMOTES GUANINE-NUCLEOTIDE EXCHANGE ON ARF1, ARF3 AND ARF6. PROMOTES THE ACTIVATION OF ARF THROUGH REPLACEMENT OF GDP WITH GTP.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: UBICITOUS.

CC -1- SIMILARITY: CONTAINS 1 SEC7 DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X99753; CAA68084.1; -

DR EMBL; U70728; AAB09591.1; -

DR PDB; 1PBV; 09-MAR-99.

DR MIM; 602488; -

DR InterPro: IPR000904; -

DR InterPro: IPR001849; -

DR Pfam; PF00169; Pf; 1.

DR Pfam; PF01369; Sec7; 1.

DR PROSITE: PS50003; PH\_DOMAIN: 1.

KW Guanine-nucleotide releasing factor; coiled coil;

KW Alternative splicing; 3D-structure.

FT DOMAIN 10 63 COILED COIL (POTENTIAL).

FT DOMAIN 72 201 SEC7.

FT DOMAIN 259 376 PH.

FT VARSPLIC 272 272 MISSING (IN ISOFORM 2).

SO SEQUENCE 400 AA; 46546 MW; 70441A58483BDOEI CRC64;

Query Match 50.0%; Score 5; DB 1; Length 400;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 APPOE 7

Db 358 APPOE 362

RESULT 40

CYH2\_MOUSE STANDARD: PRT; 400 AA.

AC P97695; O89099; 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE CYTOSKESTIN 2 (ARF NUCLEOTIDE-BINDING SITE OPENER) (ARNO PROTEIN) (CLM2) (SEC7 HOMOLOG B) (MSEC7-2).

GN PSCD2 OR SEC7B.

OS Mus musculus (Mouse), and

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090, 10116;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).

RC SPECIES=Mouse; TISSUE=Brain;

RX MEDLINE=98416066; PubMed=9744817;

RA Kim H.-S., Chen Y., Lonal P.;

RT "Complex regulation of multiple cytohesin-like genes in murine tissues and cells.";

RL FEBS Lett. 433:312-316(1998).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC SPECIES=Mouse;

RX Liu D., Zhang H., Lu J.;

RA "cDNA cloning of mouse cytohesin-2 and demonstration of its association with the integrin beta2 subunit.";

RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

RL [3]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC SPECIES=Rat;

RX PubMed=9352219;

RA Telemakos I., Benschler F., Stenlus K., Sudhof T.C., Brose N.;

RT "Rat homologues of yeast sec7p.";

RL Eur. J. Cell Biol. 74:143-149(1997).

CC -1- FUNCTION: PROMOTES GUANINE-NUCLEOTIDE EXCHANGE ON ARF1, ARF3 AND ARF6. PROMOTES THE ACTIVATION OF ARF THROUGH REPLACEMENT OF GDP WITH GTP.

CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1/CLM2-A (SHOWN HERE), 2 AND 3/CLM2-B; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: PRESENT IN ALL TISSUES TESTED, WITH HIGHEST PROTEIN LEVELS IN BRAIN AND ADRENAL.

CC -1- SIMILARITY: CONTAINS 1 SEC7 DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

Query Match 50.0%; Score 5; DB 1; Length 367;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAPTO 6  
|||||  
DB 67 VAPTO 71

RESULT 37  
GAG\_HV1W2 STANDARD; PRT: 388 AA.  
ID GAG\_HV1W2  
AC P05889;  
DT 01-NOV-1988 (rel. 09, Created)  
DT 01-FEB-1994 (rel. 28, Last sequence update)  
DT 15-DEC-1998 (rel. 37, Last annotation update)  
DE GAG POLYPROTEIN [CONTAINS: CORE PROTEINS P17, P24, P2, P7, P1, P6] (FRAGMENT).  
GN GAG.  
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).  
OC Viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11705;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86235450; PubMed=3012778;  
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,  
RA Salathuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;  
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or  
RT at risk for AIDS".  
RL Science 233:1548-1553(1986).  
CC -1- FUNCTION: PERFORMS HIGHLY COMPLEX ORCHESTRATED TASKS DURING THE  
CC ASSEMBLY, BUDDING, MATURATION, AND INFECTION STAGES OF THE VIRAL  
CC REPLICATION CYCLE. DURING VIRAL ASSEMBLY, THE PROTEINS FORM  
CC MEMBRANE ASSOCIATIONS AND SELF-ASSOCIATIONS THAT ULTIMATELY  
CC RESULT IN BUDDING OF AN IMMATURE VIRION FROM THE INFECTED CELL.  
CC GAG PRECURSORS ALSO FUNCTION DURING VIRAL ASSEMBLY TO SELECTIVELY  
CC BIND AND PACKAGE TWO PLUS STRANDS OF GENOMIC RNA.  
CC -1- PTM: THE P24 PROTEIN IS PHOSPHORYLATED.  
CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM  
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO  
CC WAS PERINATALLY INFECTED BY HER MOTHER.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC  
CC EMBL: K03457; AAB12988.1; -  
CC HSSP: P05888; IAAE.  
CC HIV: K03457; GAGSMWJ2.  
CC InterPro: IPR000071; -  
CC InterPro: IPR000721; -  
CC Pfam: PF00540; gag\_p17; 1.  
DR Pfam: PF00607; gag\_p24; 1.  
KW AIDS; Core protein; Polyprotein; Myristate; Phosphorylation;  
KW Zinc-finger.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT CHAIN 1 131 CORE PROTEIN P17 (MATRIX PROTEIN).  
FT CHAIN 132 362 CORE PROTEIN P24 (CORE ANTIGEN).  
FT CHAIN 363 376 CORE PROTEIN P2.  
FT CHAIN 377 >388 CORE PROTEIN P7 (NUCLEOCAPSID PROTEIN).  
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).  
FT NON\_TER 388 388  
SQ SEQUENCE 388 AA; 43314 MW; EF885BE10ECF7804 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 388;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 5 TOEVK 9  
|||||  
DB 309 TOEVK 313

RESULT 38  
THAB\_PAROL STANDARD; PRT: 391 AA.  
ID THAB\_PAROL  
AC Q91242;  
DT 15-JUL-1999 (rel. 38, Created)  
DT 15-JUL-1999 (rel. 38, Last sequence update)  
DT 15-JUL-1999 (rel. 38, Last annotation update)  
DE THYROID HORMONE RECEPTOR ALPHA B (THR-ALPHA-B).  
GN THRA2 OR NR1A1-B.  
OS Parachanna olivacea (Flounder).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectidae; Bothidae; Parachanna.  
OX NCBI\_Taxid=8235;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95008408; PubMed=7923940;  
RA Yamano K., Araki K., Sekikawa K., Inui Y.;  
RT "Cloning of thyroid hormone receptor genes expressed in metamorphosing  
RT flounder".  
RL Dev. Genet. 15:378-382(1994).  
CC -1- FUNCTION: HIGH AFFINITY RECEPTOR FOR TRIIODOTHYRONINE.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,  
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
CC NRI SUBFAMILY.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC  
CC EMBL: D16463; BAA03929.1; -  
CC InterPro: IPR000324; -  
CC InterPro: IPR000536; -  
CC InterPro: IPR001628; -  
CC InterPro: IPR001723; -  
CC InterPro: IPR001728; -  
DR Pfam: PF00104; hormone\_rec; 1.  
DR Pfam: PF00105; zf-C4; 1.  
DR PRINTS: PRO0047; STROIDFINGER.  
DR PRINTS: PRO0350; VITAMINDR.  
DR PRINTS: PRO0398; STROHORMNR.  
DR PRINTS: PRO0546; THYROIDHORMR.  
DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
KW Zinc-finger; Multigene family.  
FT DOMAIN 1 32 MODULATING.  
FT DNA\_BIND 33 100 NUCLEAR RECEPTOR-TYPE.  
FT ZN\_FING 33 53 CA-TYPE.  
FT ZN\_FING 71 95 CA-TYPE.  
FT DOMAIN 150 391 LIGAND-BINDING.  
SQ SEQUENCE 391 AA; 45055 MW; 19FB8631EE63D022 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 391;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 PVAPT 5  
|||||

OC Viruses; dsRNA viruses; Recoviridae; Orthoreovirus.  
 OX NCBI\_TaxID=10885;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88072071; PubMed=3686825;  
 RA Wiener J.R., Joklik W.K.;  
 RT "Comparison of the reovirus serotype 1, 2, and 3 S3 genome segments  
 encoding the nonstructural protein sigma NS.";  
 RL Virology 161:332-339(1987).  
 CC -1- FUNCTION: THIS PROTEIN HAS A POLY(C)-DEPENDENT POLY(C) POLYMERASE  
 CC ACTIVITY, AND BINDS SINGLE-STRANDED, BUT NOT DOUBLE-STRANDED, RNA.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M18390; AAA47282.1; -  
 DR PIR: B27401; MNXRT2.  
 DR InterPro: IPR002507; -  
 DR Pfam: PF01518; PolyG.pol; 1.  
 KW Core protein; RNA-binding; Transferase; RNA-directed RNA polymerase.  
 SQ SEQUENCE 366 AA; 41334 MW; EF9202C3A6FC5085 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 366;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
 |||||  
 DB 189 PVAPT 193

RESULT 35  
 RPRO\_RECVL  
 ID RPRO\_RECVL STANDARD; PRT: 366 AA.  
 AC P07940.  
 DT 01-AUG-1988 (rel. 08, Created)  
 DT 01-AUG-1988 (rel. 08, Last sequence update)  
 DT 15-JUL-1998 (rel. 36, Last annotation update)  
 DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (SIGMA NS PROTEIN) (CORE  
 DE PROTEIN) (RNA-BINDING PROTEIN).  
 GN S3.  
 OS Reovirus (type 1 / strain Lang).  
 OC Viruses; dsRNA viruses; Recoviridae; Orthoreovirus.  
 OX NCBI\_TaxID=10884;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87025802; PubMed=3767989;  
 RA George C.X., Atwater J.A., Sammel C.E.;  
 RT "Biosynthesis of reovirus-specified polypeptides. Molecular cDNA  
 RT cloning and nucleotide sequence of the reovirus serotype 1 Lang  
 RT strain S3 mRNA which encodes the nonstructural RNA-binding protein  
 RT sigma NS.";  
 RL Biochem. Biophys. Res. Commun. 139:845-851(1986).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 RA MEDLINE=88072071; PubMed=3686825;  
 RA Wiener J.R., Joklik W.K.;  
 RT "Comparison of the reovirus serotype 1, 2, and 3 S3 genome segments  
 RT encoding the nonstructural protein sigma NS.";  
 RL Virology 161:332-339(1987).  
 CC -1- FUNCTION: THIS PROTEIN HAS A POLY(C)-DEPENDENT POLY(C) POLYMERASE  
 CC ACTIVITY, AND BINDS SINGLE-STRANDED, BUT NOT DOUBLE-STRANDED, RNA.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M14325; AAA47273.1; -  
 DR EMBL: M18389; AAA47281.1; -  
 DR PIR: A25068; MNXRT2.  
 DR InterPro: IPR002507; -  
 DR Pfam: PF01518; PolyG.pol; 1.  
 KW Core protein; RNA-binding; Transferase; RNA-directed RNA polymerase.  
 FT CONFLICT 250 250 V -> I (IN REF. 2).  
 SQ SEQUENCE 366 AA; 41188 MW; B9172A6A9251CC9A CRC64;

Query Match 50.0%; Score 5; DB 1; Length 366;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
 |||||  
 DB 189 PVAPT 193

RESULT 36  
 NU1M\_PODAN  
 ID NU1M\_PODAN STANDARD; PRT: 367 AA.  
 AC P19041.  
 DT 01-NOV-1990 (rel. 16, Created)  
 DT 01-JUL-1993 (rel. 26, Last sequence update)  
 DT 15-JUL-1999 (rel. 38, Last annotation update)  
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3).  
 GN NDI.  
 OS Podospora anserina.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Podospora.  
 OX NCBI\_TaxID=5145;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S, AND A;  
 RX MEDLINE=89063443; PubMed=3197134;  
 RA Cummings D.J., Domenico J.M., Michel F.;  
 RT "DNA sequence and organization of the mitochondrial NDI gene from  
 RT Podospora anserina: analysis of alternate splice sites.";  
 RL Curr. Genet. 14:253-264(1988).  
 CC [2]  
 CC COMPLETE GENOME.  
 RP STRAIN-S;  
 RC MEDLINE=90291512; PubMed=2357736;  
 RA Cummings D.J., McNally K.L., Domenico J.M., Matsuura E.T.;  
 RT "The complete DNA sequence of the mitochondrial genome of Podospora  
 RT anserina.";  
 RL Curr. Genet. 17:375-402(1990).  
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.  
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X55026; CA38809.1; -  
 DR EMBL: X13164; CA31563.1; ALT-SEQ.  
 DR PIR: S06058; S06058.  
 DR InterPro: IPR001694; -  
 DR Pfam: PF00146; NADHdh; 1.  
 DR PROSITE: PS00667; COMPLEX1\_NDI\_1; 1.  
 DR PROSITE: PS00668; COMPLEX1\_NDI\_2; 1.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.  
 SQ SEQUENCE 367 AA; 41079 MW; 3AFC9D186EB25C2C CRC64;

QY	1	PVAPT	5
Db	284	PVAPT	288

[illegible]

Db 328 TQEVK 332

```

RESULT 33
RRPO_REV03 ID RRPO_REV03 STANDARD: PRI: 366 AA.
AC P03526;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (SIGMA NS PROTEIN) (CORE
DE PROTEIN) (RNA-BINDING PROTEIN).
CN S3.
OS Recovirus (type 3 / strain Dearing).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.
OX NCBI_TaxID=10886;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=84015379; PubMed=6312421;
RA Richardson M.A., Furuichi Y.;
RT "Nucleotide sequence of reovirus genome segment S3, encoding non-
RT structural protein sigma NS."
RL Nucleic Acids Res. 11:6399-6408(1983).
RN 12
RP SEQUENCE OF 1-17 FROM N.A.
RX MEDLINE=83017876; PubMed=6927854;
RA Antezek J.B., Cimelo R.A., Pickup D.J., Joklik W.K.;
RT "Sequence at both termini of the 10 genes of reovirus serotype 3
RT (strain Dearing).";
RL Virology 121:307-319(1982).
CC -1- FUNCTION: THIS PROTEIN HAS A POLY(C)-DEPENDENT POLY(G) POLYMERASE
CC ACTIVITY, AND BINDS SINGLE-STRANDED, BUT NOT DOUBLE-STRANDED, RNA.
CC -1- SIMILARITY: 86.3% SIMILARITY TO REOVIRUS SEROTYPE 2 SIGMA NS
CC PROTEIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC
DR DR EMBL: X01627; CAA25768.1; -.
DR DR EMBL: J02330; AAA47280.1; -.
DR DR PIR: A04126; MNKRSD.
DR DR InterPro: IPR002507; -.
DR DR Pfam: PF01518; PolyC_pol; 1.
KW Core protein; RNA-binding; Transferase; RNA-directed RNA polymerase.
SQ SEQUENCE 366 AA; 41056 MW; 201A623C4290DD7 CMC64;

Query Match 50.0%; Score 5; DB 1; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5
DB 189 PVAPT 193

RESULT 34
RRPO_REV04 ID RRPO_REV04 STANDARD: PRI: 366 AA.
AC P12002;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (SIGMA NS PROTEIN) (CORE
DE PROTEIN) (RNA-BINDING PROTEIN).
CN S3.
OS Recovirus (type 2 / strain D5/Jones).

```



```

RT      "Purification of two thermostable components of messenger
RT      ribonucleoprotein particles (mrnp) from Xenopus laevis oocytes,
RT      belonging to a novel class of RNA-binding proteins."
RL      FEBS Lett. 282:110-114(1991).
CC      -1- FUNCTION: BINDS TO CCAAT-CONTAINING 5' BOX OF THE HSP70 GENES.
CC      SEEMS TO BE A NEGATIVE REGULATORY FACTOR. ALSO BINDS TO MRNA.
CC      -1- SUBUNIT: POSSIBLY FORMS A HETERODIMER WITH P54 IN THE 6S AND 15S
CC      MRNA-BINDING PARTICLES.
CC      -1- SUBCELLULAR LOCATION: CYTOPLASMIC, EITHER FREE OR ASSOCIATED WITH
CC      RIBONUCLEOPROTEIN PARTICLES.
CC      -1- TISSUE SPECIFICITY: TESTIS AND IMMATURE OOCYTES.
CC      -1- PTM: PHOSPHORYLATION ACTIVATES IN VITRO RNA-BINDING.
CC      -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: M59454; AAA49716.1; -.
DR      PIR: B38274; B38274.
DR      HSSP: P33081; 1IMG.
DR      TRANSFAC: T00294; -.
DR      InterPro: IPR002059; -.
DR      Pfam: PF00313; CSD. 1.
DR      PRINTS: PR00050; COLDSHOCK.
DR      PROSITE: PS00352; COLD_SHOCK. 1.
DR      Transcription regulation; DNA-binding; Nuclear protein;
KW      RNA-binding; Phosphorylation.
FT      DOMAIN 44 108 CSD.
FT      CONFLICT 254 254 A -> T (IN REF. 1).
SQ      SEQUENCE 336 AA; 37202 MW; 4AD5838769C6B84D CRC64;

Query Match          50.0%; Score 5; DB 1; Length 336;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PVAPT 5
        |||||
DB      207 PVAPT 211

RESULT 30
TRPA_MAIZE          STANDARD;          PRT;          346 AA.
AC      P42390;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      TRYPTOPHAN SYNTHASE ALPHA CHAIN PRECURSOR (EC 4.2.1.20).
OS      Zea mays (Maize).
OC      Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC      Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC      Andropogoneae; Zea.
OX      NCBI_TaxID=4577;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CV. CG000237;
RX      MEDLINE=95284367; PubMed=7766899;
RA      Kramer V.C.; Kozlowski M.G.;
RT      *Structure of a maize tryptophan synthase alpha subunit gene with
RT      pth enhanced expression.
RL      Plant Mol. Biol. 27:1183-1188(1995).
CC      -1- FUNCTION: THE ALPHA SUBUNIT IS RESPONSIBLE FOR THE ALDOL CLEAVAGE
CC      OF INDOLEGLYCEROL PHOSPHATE TO INDOLE AND GLYCERALDEHYDE 3-
CC      PHOSPHATE.
CC      -1- CATALYTIC ACTIVITY: L-SERINE + 1-(INDOL-3-YL)GLYCEROL 3-PHOSPHATE
CC      = L-TRYPTOPHAN + GLYCERALDEHYDE 3-PHOSPHATE + H(2)O.
CC      -1- PATHWAY: LAST (FIFTH) STEP IN BIOSYNTHESIS OF TRYPTOPHAN.

```

```

CC      -1- SUBUNIT: Tetramer of two alpha and two beta chains.
CC      -1- SUBCELLULAR LOCATION: CHLOROPLAST (PROBABLE).
CC      -1- SIMILARITY: BELONGS TO THE TRPA FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X76713; CAA54131.1; -.
DR      HSSP: P00929; 2TSY.
DR      Makedb: 102199; -.
DR      InterPro: IPR002028; -.
DR      Pfam: PF00290; TRP-SYNTH. 1.
DR      PROSITE: PS00167; TRP-SYNTHASE ALPHA; 1.
KW      Tryptophan biosynthesis; Lyase; Chloroplast; Transl. peptide.
FT      TRANSIT 1 2 ?
FT      CHAIN 1 346 ?
SQ      SEQUENCE 346 AA; 36967 MW; 05F8FC2635071F3E CRC64;

```

```

Query Match          50.0%; Score 5; DB 1; Length 346;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 QEVKK 10
        |||||
DB      278 QEVKK 282

RESULT 31
YMA3_MYCBO          STANDARD;          PRT;          354 AA.
ID      YMA3_MYCBO
AC      Q02279;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      HYPOTHETICAL 38.1 KDA PROTEIN IN MMS 3' REGION.
OS      Mycobacterium bovis.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1765;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BCG;
RX      MEDLINE=92406887; PubMed=1527058;
RA      Mathur M.; Kolatukudy P.E.;
RT      *Molecular cloning and sequencing of the gene for mycocerosic acid
RT      synthase, a novel fatty acid elongating multifunctional enzyme, from
RT      Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guérin.
RL      J. Biol. Chem. 267:19388-19395(1992).
CC      -1- SIMILARITY: SOME, TO M.TUBERCULOSIS RV2285.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: M95808; AAA25370.1; -.
KW      Hypothetical protein.
SQ      SEQUENCE 354 AA; 37956 MW; 2C77C1259BD3686A CRC64;

```

```

Query Match          50.0%; Score 5; DB 1; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

CC -1- SUBCELLULAR LOCATION: PEROXISOMAL.
CC -1- SIMILARITY: BELONGS TO THE URICASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M10594; AAA3994.1; -.
DR EMBL: AB002807; BAA19670.1; -.
DR EMBL: AB002809; BAA19672.1; -.
DR Mendel: 16776; GYma:1087; mml6776.
DR InterPro: IPR002042; -.
DR Pfam: PF01014; URICase; 1.
DR PRINTS: PR00093; URICASE.
DR PROSITE: PS00366; URICASE; 1.
DR PROSITE: PS00342; MICROBODIES.CTER; 1.
DR Oxidoreductase; Peroxisome; Purine metabolism.
KW CONFLICT 208
KW SEQUENCE 309 AA; 35137 MW; 81B8FBF7ACAA00A9 CRC64;

Query Match
Best Local Similarity 50.0%; Score 5; DB 1; Length 309;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 QEVKK 10
DB 215 QEVKK 219

RESULT 28
ENT2_MOUSE STANDARD; PRT; 327 AA.
AC 061672;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE EQUILBRATIVE NUCLEOSIDE TRANSPORTER 2 (EQUILBRATIVE
DE NITROBENZYLMECAPTOPURINE RIBOSIDE-INSENSITIVE NUCLEOSIDE
DE (EQUILBRATIVE NEMPR-INSENSITIVE NUCLEOSIDE TRANSPORTER) (NUCLEOSIDE
DE TRANSPORTER, EI-TYPE) (36 KDA NUCLEOLAR PROTEIN HNP36) (HYDROPHOBIC
DE NUCLEOLAR PROTEIN, 36 KDA) (DELAYED-EARLY RESPONSE PROTEIN 12).
GN SLC29A2 OR ENT2 OR HNP36 OR DER12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=Fibroblast;
RX MEDLINE=95367016; PubMed=7639753;
RA Williams J.B., Lanahan A.A.;
RT "A mammalian delayed-early response gene encodes HNP36, a novel,
RT conserved nucleolar protein."
RL Biochem. Biophys. Res. Commun. 213:325-333(1995).
CC -1- FUNCTION: MEDIATES EQUILBRATIVE TRANSPORT OF PURINE AND
CC PYRIMIDINE NUCLEOSIDES, AND THE PURINE BASE HYPOXANTHINE.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR. INTEGRAL MEMBRANE
CC PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM AND A SHORT
CC FORM/HNP36 (SHOWN HERE); SEEMS TO BE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- INDUCTION: BY PLATELET DERIVED GROWTH FACTOR (PDGF) AND FIBROBLAST
CC GROWTH FACTOR (FGF).
CC -1- SIMILARITY: BELONGS TO THE SLC29A FAMILY OF TRANSPORTERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X86682; CAA60381.1; -.
DR MGD: MGI:1345278; SLC29a2.
DR InterPro: IPR002259; -.
DR Pfam: PF01733; Nucleoside tran. 1.
DR PRINTS: PRO1130; DERENTRNSPT.
KW Nuclear protein; Transmembrane; Transport; Alternative splicing.
FT TRANSMEM 2
FT TRANSMEM 22
FT TRANSMEM 32
FT TRANSMEM 52
FT TRANSMEM 63
FT TRANSMEM 83
FT TRANSMEM 159
FT TRANSMEM 179
FT TRANSMEM 194
FT TRANSMEM 214
FT TRANSMEM 231
FT TRANSMEM 251
FT TRANSMEM 267
FT TRANSMEM 287
FT TRANSMEM 303
FT TRANSMEM 323
SO SEQUENCE 327 AA; 36113 MW; 5D2D3FF4BBD59286 CRC64;

Query Match
Best Local Similarity 50.0%; Score 5; DB 1; Length 327;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APTQE 7
DB 100 APTQE 104

RESULT 29
YB56_XENLA STANDARD; PRT; 336 AA.
ID YB56_XENLA
AC P21574;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CYTOPLASMIC RNA-BINDING PROTEIN P56 (Y BOX BINDING PROTEIN-2) (Y-BOX
DE TRANSCRIPTION FACTOR) (MRNP4).
GN FRG12.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=91062413; PubMed=2247479;
RA Tafuri S.R., Wolffe A.P.;
RT "Xenopus Y-box transcription factors: molecular cloning, functional,
RT analysis and developmental regulation."
RL Proc. Natl. Acad. Sci. U.S.A. 87:9028-9032(1990).
RN [2]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC MEDLINE=92107999; PubMed=1729676;
RA Murray M.T., Schiller D.L., Franke M.W.;
RT "Sequence analysis of cytoplasmic mRNA-binding proteins of Xenopus
RT oocytes identifies a family of RNA-binding proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 89:11-15(1992).
RN [3]
RP PARTIAL SEQUENCE.
RC TISSUE-Ovary;
RX MEDLINE=92332467; PubMed=1629179;
RA Deschamps S., Viel A., Garrigos M., Denis H., le Maire M.;
RT "MRNP4, a major mRNA-binding protein from Xenopus oocytes is
RT identical to transcription factor FRG Y2."
RL J. Biol. Chem. 267:13799-13802(1992).
RN [4]
RP PARTIAL SEQUENCE.
RC TISSUE-Ovary;
RX MEDLINE=91224309; PubMed=1902800;
RA Deschamps S., Viel A., Denis H., le Maire M.;

```

```

FT CHAIN 64 255 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
FT CHAIN 256 >309 NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 309 AA; 32922 MW; 6E858EC3DDB9EA9 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5
Db 113 PVAPT 117

RESULT 26
ID URIC_SOYBN STANDARD; PRT; 309 AA.
AC P04670; P34805; P93160; P93161; O04105;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE URICASE II ISOZYME 1 (EC 1.7.3.3) (URATE OXIDASE) (MODULIN 35) (N-35)
DE (NODULE SPECIFIC URICASE).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Glycine.
OC NCBI_TaxID=3847;
OX 11
RN SEQUENCE FROM N.A.
RP STRAIN-CV. PRIZE.
RC "Suzuki H., Verma D.P.S.;
RA "Soybean nodule-specific uricase (nodulin-35) is expressed and
RT assembled into a functional tetrameric holoenzyme in Escherichia
RL coli.";
RN 12
RN SEQUENCE FROM N.A., AND SEQUENCE OF 245-262.
RP STRAIN-CV. PRIZE.
RC Nguyen T., Zelechowska M., Foster V., Bergmann H., Verma D.P.S.;
RA "Primary structure of the soybean nodulin-35 gene encoding uricase II
RT localized in the peroxisome of uninfected cells of nodules.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:5040-5044(1985).
RN 13
RN SEQUENCE FROM N.A.
RP STRAIN-CV. AKISENGOKU; TISSUE=Cotyledon, and Root nodules;
RA Takane K., Tanaka K., Tajima S., Okazaki K., Kouchi H.;
RT "Expression of uricase II gene (nodulin-35) in cotyledons of soybean
RL plants.";
RN Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN 14
RN SEQUENCE FROM N.A.
RP STRAIN-CV. AKISENGOKU; TISSUE=Root nodules;
RA Takane K., Tajima S., Kouchi H.;
RT "Two distinct uricase II (nodulin 35) genes are differentially
RL expressed in soybean plants.";
RN Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN 15
RN FUNCTION: THIS IS A SUBUNIT OF THE NODULE SPECIFIC URICASE.
CC -1- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O -> 5-HYDROXYISOURATE +
CC H(2)O(2) (5-HYDROXYISOURATE DECOMPOSE TO FORM ALLANTOIN).
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: PEROXISOMAL.
CC -1- INDUCTION: DURING NODULATION IN LEGUME ROOTS AFTER RHIZOBIUM
CC INFECTION.
CC -1- PTM: THE N-TERMINUS IS BLOCKED.

```

```

CC -1- SIMILARITY: BELONGS TO THE URICASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M63743; AAA33997.1; -.
DR EMBL: M95400; AAA34019.1; -.
DR EMBL: L00353; AAA33995.1; -.
DR EMBL: D86929; BAA1184.1; -.
DR EMBL: D86930; BAA1185.1; -.
DR EMBL: AB002810; BAA19673.1; -.
DR PIR: A25776; A25776.
DR Mendel: 13686; Glyma;1087.1.
DR Mendel: 16777; Glyma;1087;ma16777.
DR InterPro: IPR002042; -.
DR Pfam: PF01014; URICASE. 1.
DR PRINTS: PR00093; URICASE.
DR PROSITE: PS00366; URICASE. 1.
DR PROSITE: PS00342; MICROBODIES_CTER. 1.
DR Nucleotide; Oxidoreductase; Peroxisome; Purine metabolism.
FT SITE 307 309 MICROBODY TARGETING SIGNAL (POTENTIAL).
FT VARIANT 3 3 O -> K (IN CV. AKISENGOKU).
FT VARIANT 235 235 S -> T.
FT CONFLICT 60 60 D -> E (IN REF. 2).
FT CONFLICT 195 195 Y -> C (IN REF. 4).
SQ SEQUENCE 309 AA; 35138 MW; 6FB2A758FBA982 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 309;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 OEVRK 10
Db 215 OEVRK 219

RESULT 27
ID URID_SOYBN STANDARD; PRT; 309 AA.
AC C04104;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE URICASE II ISOZYME 2 (EC 1.7.3.3) (URATE OXIDASE) (MODULIN 35) (N-35)
DE (NON-SYMBIOTIC URICASE).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Glycine.
OC NCBI_TaxID=3847;
OX 11
RN SEQUENCE FROM N.A., AND SEQUENCE OF 245-262.
RP STRAIN-CV. DARE.
RC Nguyen T., Zelechowska M., Foster V., Bergmann H., Verma D.P.S.;
RA "Primary structure of the soybean nodulin-35 gene encoding uricase II
RT localized in the peroxisome of uninfected cells of nodules.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:5040-5044(1985).
RN 12
RN SEQUENCE FROM N.A.
RP STRAIN-CV. AKISENGOKU; TISSUE=Root nodules;
RA Takane K., Tajima S., Kouchi H.;
RT "Two distinct uricase II (nodulin 35) genes are differentially
RL expressed in soybean plants.";
RN Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN 15
RN FUNCTION: THIS IS A SUBUNIT OF THE NODULE SPECIFIC URICASE.
CC -1- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O -> 5-HYDROXYISOURATE +
CC H(2)O(2) (5-HYDROXYISOURATE DECOMPOSE TO FORM ALLANTOIN).
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

```

DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE HYPOTHELICAL 33.5 KDA PROTEIN IN MRPS9-YSM1 INTERGENIC REGION.  
 GN YBR147W OR YBR1124.  
 OS Saccharomyces cerevisiae (baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.  
 CC NCBI\_TaxID=4932;  
 CC [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=528C;  
 RA Enlian K.-D., Koetter P., Rose M., Becker J., Grey M., Li Z.,  
 RA Nlegmann E., Schenk-Groeninger R., Servos J., Wehner E.,  
 RA Wolter R., Brendel R., Bauer H., Braun H., Dern K., Duesterhus S.,  
 RA Grunbein R., Hedges D., Klesau P., Korol S., Krebs B., Proft M.,  
 RA Siegers K., Baur A., Boles A., Miosga T.,  
 RA Schaff-Gerstenschlaeger I., Zimmermann F.K.;  
 RL Submitted (AUG-1994) to the EMBL/Genbank/DBS databases.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: SOME, TO S.POMBE SPAC2E12.03C AND SPAC17C9.10.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC DR EMBL: Z36016; CAA85105.1; -  
 CC DR PIR: S46018; S46018.  
 CC DR SGD: S0000351; YBR147W.  
 CC KM Hypothetical protein; Transmembrane.  
 CC FT TRANSMEM 13 33 POTENTIAL.  
 CC FT TRANSMEM 45 65 POTENTIAL.  
 CC FT TRANSMEM 69 89 POTENTIAL.  
 CC FT TRANSMEM 164 184 POTENTIAL.  
 CC FT TRANSMEM 200 220 POTENTIAL.  
 CC FT TRANSMEM 239 259 POTENTIAL.  
 CC FT TRANSMEM 263 283 POTENTIAL.  
 CC SQ SEQUENCE 296 AA; 33496 MW; 50D65896A1BARB43 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 296;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 OEYKK 10  
 DB 103 OEYKK 107

RESULT 24  
 LXD2\_PHOLE STANDARD; PRT; 305 AA.  
 AC Q06878;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE ACT1 TRANSFERASE (EC 2.3.1.-) (ACT1) (MYRISTOYL-ACP-SPECIFIC  
 DE THIOESTERASE).  
 GN LUXD.  
 OS Photobacterium leiognathl.  
 CC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;  
 CC Photobacterium.  
 CC NCBI\_TaxID=658;  
 CC [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=741;  
 RA MEDLINE=93331531; PubMed=8472957;  
 RA Chao Y.-F., Weng S.-F., Lin J.-W.;  
 RT "Sequence of the luxD gene encoding acyltransferase of the lux operon  
 RT from Photobacterium leiognathl.";  
 RT Gene 126:155-156(1993).

CC -1- FUNCTION: ACYL TRANSFERASE IS PART OF THE FATTY ACID REDUCTASE  
 CC SYSTEM REQUIRED FOR ALDEHYDE BIOSYNTHESIS. IT PRODUCES FATTY  
 CC ACIDS FOR THE LUMINESCENT REACTION.  
 CC -1- PATHWAY: FIRST STEP IN THE BIOLUMINESCENT FATTY ACID REDUCTASE  
 CC COMPLEX.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC DR EMBL: X65612; CAA46563.1; -  
 CC DR PIR: JN0518; JN0518.  
 CC DR HSP: P05321; ITHT.  
 CC KM Luminescence; Transferase; Acyltransferase.  
 CC SQ SEQUENCE 305 AA; 34384 MW; 0BDAA76D08913164 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 OEYKK 10  
 DB 219 OEYKK 223

RESULT 25  
 POLG\_HCVH7 STANDARD; PRT; 309 AA.  
 AC P27955;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE GENOME POLYPROTEIN (CONTAINS: MATRIX PROTEIN (ENVELOPE PROTEIN M);  
 DE MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL PROTEIN NS1) (FRAGMENT).  
 DE Hepatitis C virus (isolate HCV7) (HCV).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 CC Hepacivirus.  
 CC NCBI\_TaxID=11109;  
 CC [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=91112009; PubMed=1846505;  
 RA Weiner A.J., Brauer M.J., Rosenblatt J., Richman K.H., Tung J.,  
 RA Crawford K., Bonino F., Saracco G., Choo Q.-L., Houghton M.,  
 RA Han J.H.;  
 RT "Variable and hypervariable domains are found in the regions of HCV  
 RT corresponding to the flavivirus envelope and NS1 proteins and the  
 RT pestivirus envelope glycoproteins";  
 RT Virology 180:842-848(1991).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC DR EMBL: X53133; CAA37293.1; -  
 CC DR InterPro: IPR002519; -  
 CC DR InterPro: IPR002521; -  
 CC DR InterPro: IPR002531; -  
 CC DR Pfam: PF01560; HCV\_NSI; 1.  
 CC DR Pfam: PF01542; HCV\_Core; 1.  
 CC DR Pfam: PF01539; HCV\_Env; 1.  
 CC KW Polyprotein; Glycoprotein; Coat protein; Envelope protein;  
 KW Transmembrane; Nonstructural protein.  
 FT NON\_TER 1 63  
 FT CHAIN <1 63 MATRIX PROTEIN (POTENTIAL).

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: U34816; AAC43647.1; -  
 DR EMBL: AE000017; AAB95803.1; -  
 KM Hypothetical protein.  
 SQ SEQUENCE 250 AA; 28662 MW; 4378A2834BBB7877 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 83;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 OEYK 10  
 DB 24 OEYK 28

RESULT 21  
 CB24\_ARATH STANDARD; PRT; 251 AA.  
 ID CS24\_ARATH  
 AC P27521;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CHLOROPHYLL A-B BINDING PROTEIN 4 PRECURSOR (LHCI TYPE III CAB-4)  
 DE (LHCP).  
 DE CAB4 OR P1P2.20.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OC NCBI\_Taxid=3702;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang H., Hanley S., Goodman H.M.;  
 RT "Isolation, characterization and chromosomal location of a new cab  
 RL gene from Arabidopsis thaliana";  
 RL Plant Physiol. 96:1387-1388(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-CV. COLUMBIA;  
 RA Chloé N., Robert C., Brothier P., Mincker P., Catolico L.,  
 RA Chloé N., Saurin W., Melsenbach J., Mewes H.-W., Lemcke K.,  
 RA Mayer K.F.X., Queller F., Salanoubat M.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: THE LIGHT-HARVESTING COMPLEX (LHC) FUNCTIONS AS A LIGHT  
 CC RECEPTOR. IT CAPTURES & DELIVERS EXCITATION ENERGY TO PHOTOSYSTEMS  
 CC WITH WHICH IT IS CLOSELY ASSOCIATED. THE N-TERMINUS OF THE PROTEIN  
 CC EXTENDS INTO THE STROMA WHERE IT IS INVOLVED WITH ADHESION OF  
 CC GRANAL MEMBRANES AND PHOTOREGULATED BY REVERSIBLE PHOSPHORYLATION  
 CC OF ITS THREONINE RESIDUES; BOTH ARE BELIEVED TO MEDIATE THE  
 CC DISTRIBUTION OF EXCITATION ENERGY BETWEEN PHOTOSYSTEMS I AND II.  
 CC -1- SUBUNIT: THE LHC COMPLEX CONSISTS OF CHLOROPHYLLS (A & B) AND  
 CC CHLOROPHYLL A-B BINDING PROTEINS.  
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: M63931; AAA32760.1; -  
 DR EMBL: AL132955; CAB61973.1; -  
 DR InterPro: IPR001344; -  
 DR Pfam: PF00504; chloroa\_b-bind.1.  
 KM Chlorophyll; Photosynthesis; Photosystem I; Photosystem II;  
 KM Thylakoid membrane; Chloroplast; Transil peptide; Multigene family;

KM Transmembrane; Phosphorylation.  
 FT TRANST 1 ? CHLOROPLAST (POTENTIAL).  
 FT CHAIN 2 251 CHLOROPHYLL A-B BINDING PROTEIN 4.  
 FT TRANSMEM 211 227 POTENTIAL.  
 SQ SEQUENCE 251 AA; 27733 MW; DCB8390C2AED9D22 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 83;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTQE 7  
 DB 196 APTQE 200

RESULT 22  
 URED\_SYNY3 STANDARD; PRT; 270 AA.  
 ID URED\_SYNY3  
 AC P73047;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE UREASE ACCESSORY PROTEIN URED.  
 DE URED OR SLI1639.  
 OS Synechocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
 OC NCBI\_Taxid=1148;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; Pubmed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Ntaro K., Okumura S.,  
 RA Shimp S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
 RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions";  
 RL DNA Res. 3:109-136(1996)  
 CC -1- FUNCTION: PROBABLY FACILITATES NICKEL INCORPORATION.  
 CC -1- SIMILARITY: BELONGS TO THE URED FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: D90903; BAA17068.1; ALT\_INIT.  
 DR InterPro: IPR002669; -  
 DR Pfam: PF01774; Ured; 1.  
 KM Nickel.  
 SQ SEQUENCE 270 AA; 30522 MW; FAA820E257B69286 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 270;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TOEYK 9  
 DB 254 TOEYK 258

RESULT 23  
 YB27\_YEAST STANDARD; PRT; 296 AA.  
 ID YB27\_YEAST  
 AC P38278;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)

RP SEQUENCE.  
 RX MEDLINE=88000549; PubMed=3307909;  
 RA Dunbar J.C., Bradshaw R.A.;  
 RT "Amino acid sequence of guinea pig prostate kallikrein.";  
 RL Biochemistry 26:3471-3478(1987).  
 CC -1- FUNCTION: GLANDULAR KALLIKEINS CLEAVE MET-LYS AND ARG-SER BONDS  
 CC IN KININOGEN TO RELEASE LYS-BRADYKININ.  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-L-XAA BONDS IN  
 CC SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
 CC KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
 CC MET-L-XAA OR LEU-L-XAA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPsin FAMILY. KALLIKREIN SUBFAMILY.  
 DR PIR: A27207; A27207.  
 DR HSSP: P00757; 1SGF.  
 DR MEROPS: S01.160; -.  
 DR InterPro: IPR001254; -.  
 DR InterPro: IPR001314; -.  
 DR Pfam: PF00089; TRYPSIN\_1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PROSITE: PS00134; TRYPSIN\_HIS\_1.  
 DR PROSITE: PS00135; TRYPSIN\_SER\_1.  
 KM Hydrolyse; Serine protease; kininogenase; glycoprotein.  
 FT ACT\_SITE 41 41 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 96 96 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 191 191 CHARGE RELAY SYSTEM.  
 FT CARBOHD 78 78 N-LINKED (GLCNAC. . .).  
 FT CARBOHD 169 169 N-LINKED (GLCNAC. . .).  
 FT DISULFD 7 151 BY SIMILARITY.  
 FT DISULFD 26 42 BY SIMILARITY.  
 FT DISULFD 128 197 BY SIMILARITY.  
 FT DISULFD 162 176 BY SIMILARITY.  
 FT DISULFD 187 212 BY SIMILARITY.  
 FT VARIAM 50 50 K -> W.  
 SO SEQUENCE 239 AA; 25969 MW; 56DC81BC10D49A64 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PROEV 8  
 DB 118 PROEV 122  
 RESULT 19  
 Y984\_CAMJE STANDARD; PRT; 246 AA.  
 AC P45491; G9PNV5;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HYPOTHETICAL PROTEIN CJO984.  
 GN CJO984.  
 OS Campylobacter jejuni.  
 CC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 CC Campylobacter.  
 CC NCBI\_TaxID=197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 43431;  
 RX MEDLINE=95247673; PubMed=7730270;  
 RA Hant E.K., Chan V.L.;  
 RT "Expression and characterization of Campylobacter jejuni  
 RT benzoylglycine amidohydrolase (Hipuricase) gene in Escherichia  
 RT coli.";  
 RL J. Bacteriol. 177:2396-2402(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCIC 11168;  
 RX MEDLINE=20150912; PubMed=10688204;  
 RA Parthill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,

RA Basham D., Chillingworth T., Davies R.M., Felwell T., Holtroyd S.,  
 RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
 RA Whitehead S., Barrett B.G.;  
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
 RT reveals hypervariable sequences.";  
 RL Nature 403:665-668(2000).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch))  
 CC -----  
 DR EMBL: Z36940; CAA85397.1; -.  
 DR EMBL: AL139076; CAB73240.1; -.  
 KM Hypothetical protein.  
 FT CONFLICT 16 16 V -> T (IN REF. 1).  
 FT CONFLICT 21 21 N -> D (IN REF. 1).  
 FT CONFLICT 42 42 E -> G (IN REF. 1).  
 FT CONFLICT 51 51 Q -> K (IN REF. 1).  
 FT CONFLICT 118 118 I -> V (IN REF. 1).  
 FT CONFLICT 160 160 E -> K (IN REF. 1).  
 FT CONFLICT 212 212 V -> I (IN REF. 1).  
 SO SEQUENCE 246 AA; 28820 MW; 378876C1703D70CF CRC64;

Query Match 50.0%; Score 5; DB 1; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 82;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APTOE 7  
 DB 75 APTOE 79  
 RESULT 20  
 Y687\_MYCPN STANDARD; PRT; 250 AA.  
 AC Q50315;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HYPOTHETICAL PROTEIN MPN687 (K05\_ORF250).  
 GN MPN687 OR MP155.  
 OS Mycoplasma pneumoniae.  
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 CC Mycoplasmataceae; Mycoplasma.  
 CC NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=96177562; PubMed=8604303;  
 RA Hilbert H., Himmelreich R., Plagens H., Herrmann R.;  
 RT "Sequence analysis of 56 kb from the genome of the bacterium  
 RT Mycoplasma pneumoniae comprising the dna region, the atp operon and a  
 RT cluster of ribosomal protein genes.";  
 RL Nucleic Acids Res. 24:628-639(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633;  
 RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

```

RL Nucleic Acids Res. 18:3069-3069(1990).
RN [2]
RP SEQUENCE OF 203-217 FROM N.A.
RC STRAIN-TAK:
RX MEDLINE-89345113; PubMed-2762136;
RA Plenlazez N.J., Velarde J. Jr., Plenlazez D., Luftig R.B.;
RT "Nucleotide sequence of human enteric adenovirus type 41 hexon-
associated protein VIII precursor (pVIII) including the early region
E3 promoter."
RL Nucleic Acids Res. 17:5398-5398(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X52532; CAB38632.1; -.
DR EMBL: X15137; CAA33236.1; -.
DR PIR: S04851; S04851.
DR PIR: S10212; S10212.
KM Late protein; Phosphorylation.
SQ SEQUENCE 217 AA; 24740 MW; B6ECDAB63C24EC3E CRC64;

Query Match 50.0%; Score 5; DB 1; Length 217;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5
Db 118 PVAPT 122

RESULT 16
ID YD67_SCHPO STANDARD; PRT; 217 AA.
AC 010319;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 24.9 KDA PROTEIN C17G8.07 IN CHROMOSOME 1.
GN SPAC17G8.07.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
CC Schizosaccharomycetales; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO HUMAN AFG AND ENL AND YEAST TRANSCRIPTION
CC INITIATION FACTOR TFIIF SMALL SUBUNIT (TFG3/NC1).
CC -1- SIMILARITY: STRONG. TO YEAST YNL107M.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z69795; CAA33690.1; -.
KM Hypothetical protein.
SQ SEQUENCE 217 AA; 24946 MW; A274C2AFB0620D7F CRC64;

Query Match 50.0%; Score 5; DB 1; Length 217;

```

```

Best Local Similarity 100.0%; Pred. No. 74;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QEVKK 10
Db 191 QEVKK 195;

RESULT 17
ID TRKA_PYRHO STANDARD; PRT; 228 AA.
AC 057719;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TRK SYSTEM POTASSIUM UPTAKE PROTEIN TRKA HOMOLOG.
GN TRKA OR PH1984.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Pyrococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-073;
RX MEDLINE-98344137; PubMed-9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kusuda N., Oguchi A.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
CC -1- FUNCTION: PART OF A POTASSIUM TRANSPORT SYSTEM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TRKA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000007; BAA3111.1; -.
KM Transport: Potassium transport; NAD.
FT NP_BIND 2 NAD (POTENTIAL).
SQ SEQUENCE 228 AA; 25323 MW; F64E2D14EFB3FFD6 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 228;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QEVKK 10
Db 215 QEVKK 219

RESULT 18
ID KUK2_CAVPO STANDARD; PRT; 239 AA.
AC P12323;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GRANDULAR KALLIKREIN, PROSTATIC (EC 3.4.21.35) (TISSUE KALLIKREIN)
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]

```

Query Match

50.0%; Score 5; DB 1; Length 217;

Query Match

50.0%; Score 5; DB 1; Length 228;

```

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M13821; AAA8544.1; -
CC PIR: A24499; IMPSB; -
CC SEQUENCE 201 AA; 22087 MW; 25C0694D288FDAAB CRC64;
DR
SQ
Query Match
Best Local Similarity 100.0%; Score 5; DB 1; Length 201;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 PROEV 8
DB 156 PROEV 160
RESULT 13
VAD1_TREPA STANDARD; PRT; 206 AA.
ID VAD1_TREPA
AC 083443;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE V-TYPE ATP SYNTHASE SUBUNIT D 1 (EC 3.6.1.34) (V-TYPE ATPASE SUBUNIT D
DE 1).
GN ATPD1 OR TP0428.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
CC NCBI_TaxID=160;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Childsbarrow M., Uitterlind T.,
RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete."
RL Science 281:375-388(1998).
CC -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC GRADIENT ACROSS THE MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE V-ATPASE D SUBUNIT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE001220; AAC65414.1; -
CC TIGR: TP0428; -
DR InterPro: IPR002699; -
DR Pfam: PF01813; ATP-synt_D; 1.
KM Hydrolyase: ATP synthetase; Hydrogen ion transport.
SQ SEQUENCE 206 AA; 23803 MW; BC33C36FC5795483 CRC64;
Query Match
Best Local Similarity 100.0%; Score 5; DB 1; Length 206;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 6 OEYKK 10
DB 80 OEYKK 84
RESULT 14
VG40_BPPH8 STANDARD; PRT; 209 AA.
ID VG40_BPPH8
AC P14817; P14818;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PROTEINS 40A/40B.
GN 40.
OS Bacteriophage phi-80.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC Lambda phage group.
OX NCBI_TaxID=10713;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89011978; PubMed=3172225;
RA Ogawa T., Ogawa H., Tomizawa J.;
RT "Organization of the early region of bacteriophage phi 80. Genes and
RT proteins."
RL J. Mol. Biol. 202:537-550(1988).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X13065; CAA31470.1; -
DR EMBL: X13065; CAA31469.1; -
DR PIR: S04827; S04827.
DR PIR: S04826; S04826.
KM Early protein.
FT CHAIN 1
FT 115
SQ SEQUENCE 209 AA; 22974 MW; DD399BA02E7643D CRC64;
Query Match
Best Local Similarity 100.0%; Score 5; DB 1; Length 209;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 5 TOEVK 9
DB 59 TOEVK 63
RESULT 15
V33P_ADEA1 STANDARD; PRT; 217 AA.
ID V33P_ADEA1
AC P19416; G64818;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE 33 KDA PHOSPHOPROTEIN.
OS Human adenovirus type 41.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10524;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TAK;
RX MEDLINE=90272433; PubMed=2349115;
RA Siemenda S.B., Pieniazek N.J., Velarde J. Jr., Pieniazek D.,
RA Lufsig R.B.;
RT "Nucleotide sequence of the region coding for 100K and 33K proteins
RT of human enteric adenovirus type 41 (Tak).";

```



```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE001604: AAD18337.1; -.
DR EMBL: AE002217: AAF8402.1; -.
DR EMBL: AP002545: BAA98394.1; -.
DR TIGR: CP0584; -.
DR InterPro: IPR001059; -.
DR Pfam: PF01132; EFP: 1.
DR PROSITE: PS01275; EFP: 1.
KM Protein biosynthesis; Elongation factor.
SQ SEQUENCE 185 AA; 20673 MW; 6C6432CFB70354A CRC64;

Query Match
Best Local Similarity 50.0%; Score 5; DB 1; Length 185;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QYEVK 9
Db 57 QYEVK 61

RESULT 10
VGG_BPPBK STANDARD; PRT; 187 AA.
ID_VGG_BPPBK
AC Q38042;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE MAJOR SPIKE PROTEIN (G PROTEIN) (GPG).
G.
OS Bacteriophage phi-K.
OC Viruses; dsDNA viruses; Microviridae; Microvirus.
OX NCBI_TaxID=10848;
RN [1]
RP SEQUENCE FROM N.A.
RA Kodaira K.-I., Tadokoro H., Suzuki K., Oki M., Manto S., Taketo A.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR SPIKE COMPONENT. INVOLVED IN THE ATTACHMENT TO
CC THE BACTERIAL HOST.
CC -1- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE P, G, AND
CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X60323: CAA42892.1; -.
DR HSSP: P03643; IAL0.
KM Coat protein.
SQ SEQUENCE 187 AA; 19569 MW; C5C5CEC891241701 CRC64;

Query Match
Best Local Similarity 50.0%; Score 5; DB 1; Length 187;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5
Db 26 PVAPT 30

RESULT 11

```

```

RL29_MYCGE
ID RL29_MYCGE STANDARD; PRT; 200 AA.
AC P47405;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L29.
GN RPMC OR RPL29 OR MG159.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
CC -1- SIMILARITY: BELONGS TO THE L29P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U39696; AAC71377.1; -.
DR TIGR: MG159; -.
DR InterPro: IPR001854; -.
DR Pfam: PF00831; Ribosomal_L29; 1.
DR PROSITE: PS00579; RIBOSOMAL_L29; 1.
KM Ribosomal protein.
SQ SEQUENCE 200 AA; 23258 MW; 5BC115E52A5EA92 CRC64;

Query Match
Best Local Similarity 50.0%; Score 5; DB 1; Length 200;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QEVKK 10
Db 170 QEVKK 174

RESULT 12
IMWU_BPSPB STANDARD; PRT; 201 AA.
ID IMWU_BPSPB
AC P06650;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE IMMUNITY PROTEIN.
GN D.
OS Bacteriophage SP-beta.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC Lambda phage group.
OX NCBI_TaxID=10722;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2;
RX MEDLINE=86304188; PubMed=3091583;
RA McLaughlin J.R., Wong H.C., Ting Y.E., van Arsdell J.N., Chang S.;
RT "Control of lysogeny and immunity of Bacillus subtilis temperate
RL Bacteriophage SP-beta by its d gene.";

```

RT "Cloning, expression, and sequence conservation of  
 RT pathogenesis-related gene transcripts of potato.";   
 RL Mol. Plant Microbe Interact. 2:325-331(1989).  
 CC -I- INDUCTION: BY WOUNDING AND ELICITOR TREATMENTS  
 CC -I- SIMILARITY: BELONGS TO THE BETV1 FAMILY OF PATHOGENESIS-RELATED  
 CC PROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M25156; AAA03020.1; -;  
 DR EMBL; M29042; AAA02829.1; -;  
 DR PIR; S11869; S11869.  
 DR HSSP; P15494; IRTV.  
 DR InterPro; IPR000916; -;  
 DR Pfam; PF00407; Bet\_v.I; 1.  
 DR PRINTS; PR00634; BETALLERGEN.  
 DR PROSITE; PS00451; PATHOGENESIS\_BETV1; 1.  
 DR Pathogenesis-related protein; Multigene family.  
 KW SEQUENCE 155 AA; 17200 MW; 02A78E6D26A7CD9 CRC64;  
 SO  
 Query Match 50.0%; Score 5; DB 1; Length 155;  
 Best Local Similarity 100.0%; Pred. No. 56;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 PVAPT 5  
 |||||  
 Db 13 PVAPT 17  
 RESULT 8  
 ID RPL1\_BPMU STANDARD; PRT; 174 AA.  
 AC P06019;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE REPRESSOR PROTEIN Cl.  
 GI OR 1.  
 GN Bacteriophage Mu.  
 OS Bacteriophage Mu.  
 OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.  
 OX NCBI\_TaxID=10677;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=83012203; PubMed=6214696;  
 RA Priess H., Kamp D., Kahmann R., Brauer B., Delius H.;  
 RT "Nucleotide sequence of the immunity region of bacteriophage Mu.";   
 RL Mol. Genet. 186:315-321(1982).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Priess H., Brauer B., Schmidt C., Kamp D.;  
 RT "Sequence of the left end of Mu.";   
 RL (in) Symonds N., Toussaint A., van de Putte P., Howe M.M. (eds.);  
 RL Phage Mu, pp.277-296, Cold Spring Harbor Laboratory Press,  
 RL New York (1987).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Morgan G., Hattuli G., Hendrix R.;  
 RT "Genome of bacteriophage Mu and comparison with the Haemophilus  
 RT influenzae Mu-like prophage Plumu.";   
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; V01464; CA24711.1; -;  
 DR EMBL; M64097; AAA32376.1; -;  
 DR EMBL; AF083977; AAF01132.1; -;  
 DR PIR; S07291; S07291.  
 DR HSSP; P07636; ITNS.  
 KW Transcription regulation; Repressor; DNA-binding.  
 SO SEQUENCE 174 AA; 19213 MW; 11A4B3F5B5E49C5 CRC64;  
 Oy 6 QEVKK 10  
 |||||  
 Db 168 QEVKK 172  
 RESULT 9  
 ID EFPL\_CHLPN STANDARD; PRT; 185 AA.  
 AC Q92900; Q92906;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE ELONGATION FACTOR P 1 (EF-P 1).  
 GN EFPL OR CPN0184 OR CP0584.  
 OS Chlamydia pneumoniae (Chlamydiophila pneumoniae).  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydiophila.  
 OX NCBI\_TaxID=83558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CWL029;  
 RX MEDLINE=99206606; PubMed=10192388;  
 RA Kalman S., Mitchell W., Marathe R., Lamme C., Fan J., Hyman R.W.,  
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";   
 RL Nat. Genet. 21:385-389(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AR39;  
 RX MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunning R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Umayam L.A., Utterback T.,  
 RA Berry K., Bass S., Linher K., Weidman J., Kouri H., Craven B.,  
 RA Bowman C., Dodson R., Gwin M., Nelson W., Deboy R., Kolonay J.,  
 RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;  
 RT "Genome sequences of Chlamydia trachomatis Morn and Chlamydia  
 RT pneumoniae AR39.";   
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J138;  
 RX MEDLINE=20330349; PubMed=10871362;  
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
 RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
 RT from Japan and CWL029 from USA.";   
 RL Nucleic Acids Res. 28:2311-2314(2000).  
 CC -----  
 CC -I- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT  
 CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED  
 CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING  
 CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING  
 CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE  
 CC (BY SIMILARITY)  
 CC -----  
 CC -I- PATHWAY: PROTEIN BIOSYNTHESIS.  
 CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -I- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

Db 1099 TOEVKK 1104

RESULT 5  
HCC2\_CRYCO STANDARD; PRT: 102 AA.  
ID HCC2\_CRYCO  
AC 001238;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-JUL-1993 (Rel. 26, Last annotation update)  
DE MAJOR BASIC NUCLEAR PROTEIN 2 (P14 PROTEIN ALPHA CHAIN).  
CN HCC2.  
OS Cryptocodinium cohnii (Dinoflagellate).  
OC Eukaryota; Alveolata; Dinophyceae; Gonyaulacales; Cryptocodiniaceae;  
OC Cryptocodinium.  
OX NCBI\_Taxid=2866;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HND;  
RX MEDLINE=92111321; PubMed=1764969;  
RA Sala-Rovira M., Gerard M.L., Caput D., Jacques F.,  
RA Sayer-Gobillard M.O., Vernet G., Herzog M.;  
RT "Molecular cloning and immunolocalization of two variants of the  
RT major basic nuclear protein (HCC) from the histone-less eukaryote  
RT Cryptocodinium cohnii (Pyrrhophyta).";  
RL Chromosoma 100:510-518(1991).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X58444; CAA1350.1; -.  
DR PIR: S14644; S14644.  
DR PIR: S14640; S14640.  
KM Nuclear protein; DNA-binding.  
SQ SEQUENCE 102 AA; 10982 MW; 28D4C23E72C0CF28 CRC64;

Query Match 50.0%; Score 5; DB 1; Length 102;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QEVKK 10  
Db 43 QEVKK 47

RESULT 6  
KACA\_RABIT STANDARD; PRT: 103 AA.  
ID KACA\_RABIT  
AC P01840;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG KAPPA-B4 CHAIN C REGION.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_Taxid=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=8330036; PubMed=6412231;  
RA Emorine L., Dreher K.L., Kind T.J., Max E.E.;  
RT "Rabbit immunoglobulin kappa genes: structure of a germline b4  
RT allele type J-C locus and evidence for several b4-related sequences in  
RT the rabbit genome";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:5709-5713(1983).  
RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=82060334; PubMed=6795636;  
RA Heidmann O., Auffray C., Cazenave P.-A., Rougeon F.;  
RT "Nucleotide sequence of constant and 3' untranslated regions of a  
RT kappa immunoglobulin light chain mRNA of a homozygous b4 rabbit";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:5802-5806(1981).  
RN [3]  
RP SEQUENCE.  
RX MEDLINE=75133568; PubMed=1091650;  
RA Chen K.C.S., Kind T.J., Krause R.M.;  
RT "Primary structure of the L chain from a rabbit homogeneous antibody  
RT to streptococcal carbohydrate. II. Sequence determination of peptides  
RT from tryptic and peptic digests";  
RL J. Biol. Chem. 250:3289-3296(1975).  
CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM ANTIBODY TO THE  
CC SPECIFIC CARBOHYDRATE OF GROUP C STREPTOCOCCI & WAS ISOLATED FROM  
CC THE SERUM OF A SINGLE RABBIT.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X00231; CAA25051.1; -.  
DR PIR: A02122; K4RB.  
DR InterPro: IPR000495; -.  
DR InterPro: IPR003006; -.  
DR Pfam: PF00047; 19; 1.  
DR PROSITE: PS00290; IG\_MHC; FALSE\_NEG.  
KM Immunoglobulin domain; Immunoglobulin C region.  
FT NON\_TER 1 85  
FT DISULFID 26 85  
FT DISULFID 103 103  
FT CONFLICT 58 58  
SQ SEQUENCE 103 AA; 11043 MW; 5FC5AC8B60E68DBD CRC64;

Query Match 50.0%; Score 5; DB 1; Length 103;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5  
Db 2 PVAPT 6

RESULT 7  
PRSL\_SOLITU STANDARD; PRT: 155 AA.  
ID PRSL\_SOLITU  
AC P17641;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-MAY-1992 (Rel. 22, Last annotation update)  
DE PATHOGENESIS-RELATED PROTEIN STH-21.  
GN STH-21.  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
OC Solanales; Solanaceae; Solanum.  
OX NCBI\_Taxid=4113;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91346678; PubMed=2102864;  
RA Matton D.P., Bell B., Brisson N.;  
RT "Nucleotide sequence of a pathogenesis-related gene of potato";  
RL Plant Mol. Biol. 14:863-865(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92190624; PubMed=2520162;  
RA Matton D.P., Brisson N.;

Query Match 60.0%; Score 6; DB 1; Length 617;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVKK 10  
|||||  
DB 412 TOEVKK 417

## RESULT 3

ENV\_FOAMY STANDARD; PRT; 985 AA.

ID ENV\_FOAMY  
AC P14351;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE ENV POLYPROTEIN (COAT POLYPROTEIN).  
GN ENV.  
OS Human spumaretrovirus (Foamy virus).  
OC Viruses; Retroid viruses; Retroviridae; Spumaviruses.  
OX NCBI\_TaxID=11963;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88004420; PubMed=2820721;  
RA Fluegel R.M., Rethwilm A., Maurer B., Darai G.:  
RT "Nucleotide sequence analysis of the env gene and its flanking  
RL EMD J. 6:2077-2084(1987).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
CC -----  
DR EMBL: X05591; CAA29086.1; -  
DR EMBL: M54578; AAA46123.1; -  
DR PIR: C29685; VCLJSP.  
KW Coat protein; Transmembrane; Polypeptide; Glycoprotein.  
FT TRANSMEM 64 84  
FT TRANSMEM 958 978  
FT CARBOHYD 21 21  
FT CARBOHYD 105 105  
FT CARBOHYD 137 137  
FT CARBOHYD 179 179  
FT CARBOHYD 282 282  
FT CARBOHYD 307 307  
FT CARBOHYD 342 342  
FT CARBOHYD 387 387  
FT CARBOHYD 401 401  
FT CARBOHYD 419 419  
FT CARBOHYD 524 524  
FT CARBOHYD 553 553  
FT CARBOHYD 779 779  
FT CARBOHYD 805 805  
FT CARBOHYD 830 830  
SQ SEQUENCE 985 AA; 113494 MW; 60F588EA931BE0AF CRC64;

Query Match 60.0%; Score 6; DB 1; Length 985;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVKK 10  
|||||  
DB 181 TOEVKK 186

RESULT 4

ATX9\_TETTH STANDARD; PRT; 1133 AA.

AC ATX9\_TETTH  
AC 095050;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROBABLE CATION-TRANSPORTING ATPASE 9 (EC 3.6.1.-).  
GN TP9A.  
OS Tetrahymena thermophila.  
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;  
OC Tetrahymena; Tetrahymena.  
OX NCBI\_TaxID=5911;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Cu428;  
RA Wang S., Takeyasu K.:  
RT "The starvation-induced P-type ATPase in Tetrahymena thermophila."  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O - ADP + ORTHOPHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2  
CC ATPASES). SUBFAMILY V.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
CC -----  
DR EMBL: U41063; AAB08071.1; -  
DR InterPro: IPR001757; -  
DR Pfam: PF00122; E1-E2\_ATPase; 4.  
DR PROSITE: PS00154; ATPASE\_E1-E2; 1.  
KW Hydrolyase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.  
FT DOMAIN 1 6  
FT TRANSMEM 7 28  
FT TRANSMEM 29 34  
FT TRANSMEM 35 53  
FT TRANSMEM 54 167  
FT TRANSMEM 168 190  
FT TRANSMEM 191 193  
FT TRANSMEM 194 212  
FT TRANSMEM 213 363  
FT TRANSMEM 364 383  
FT TRANSMEM 384 396  
FT TRANSMEM 397 418  
FT TRANSMEM 419 887  
FT TRANSMEM 888 906  
FT TRANSMEM 907 915  
FT TRANSMEM 916 931  
FT TRANSMEM 932 948  
FT TRANSMEM 949 972  
FT TRANSMEM 973 994  
FT TRANSMEM 995 1018  
FT TRANSMEM 1019 1030  
FT TRANSMEM 1031 1050  
FT TRANSMEM 1051 1101  
FT TRANSMEM 1102 1124  
FT TRANSMEM 1125 1133  
FT MOD. RES 451 451  
FT METAL 827 827  
FT METAL 831 831  
SQ SEQUENCE 1133 AA; 129669 MW; 6875FFB651BA027 CRC64;

Query Match 60.0%; Score 6; DB 1; Length 1133;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVKK 10  
|||||



837	4	40.0	304	1	HEX1_HUMAN	Q9YJ33 homo sapien	910	4	40.0	323	1	MCH_METUA	Q9S030 methanococc
838	4	40.0	304	1	META_THEMEA	Q9YJ33 thermotoga	911	4	40.0	323	1	OTX1_BRARE	Q91994 brachydanio
839	4	40.0	304	1	YONI_CAEBL	Q09522 caenorhabdi	912	4	40.0	323	1	PF27_MROSE	P52875 mus musculu
840	4	40.0	305	1	BLAC_STRLA	P35193 streptomyce	913	4	40.0	323	1	YH8_MYCTU	O50668 mycobacteri
841	4	40.0	305	1	DDL8_ECOLI	P07862 escherichia	914	4	40.0	324	1	YH8_MYCTU	P40880 hordeum vul
842	4	40.0	305	1	FBR1_SCHPO	P35551 schizosacch	915	4	40.0	324	1	CAHC_HORVU	O06115 lactobacilli
843	4	40.0	305	1	MALM_SALTY	P26478 salmoneilla	916	4	40.0	324	1	PTGA_BACST	P42015 bacillus st
844	4	40.0	305	1	MALM_SALTY	P26478 salmoneilla	917	4	40.0	324	1	RADA_SULSO	O55075 sulfobius
845	4	40.0	305	1	ROAO_HUMAN	O61151 homo sapien	918	4	40.0	324	1	YC30_CVAPA	P48271 cyanophora
846	4	40.0	305	1	YORK_BACSU	P45927 bacillus su	919	4	40.0	325	1	MODD_MYCBO	O30620 mycobacteri
847	4	40.0	306	1	YOPD_YEREN	P37132 yeastina en	920	4	40.0	325	1	MODD_MYCBO	O30906 mycobacteri
848	4	40.0	307	1	ALF_HELPY	Q9ZM96 helicobacte	921	4	40.0	325	1	WZB8_ECOLI	P33272 escherichia
849	4	40.0	307	1	ALF_HELPY	Q9ZM96 helicobacte	922	4	40.0	325	1	WZB8_ECOLI	O33953 escherichia
850	4	40.0	307	1	CRTB_SYNP7	P37269 synecococc	923	4	40.0	325	1	WZB8_SHIFL	P95730 shigella dy
851	4	40.0	307	1	GP7D_CHLPS	O46257 chlamydia p	924	4	40.0	325	1	WZB8_SHIFL	P37792 shigella fl
852	4	40.0	307	1	VP53_BPAPS	O921P5 bacterioph	925	4	40.0	326	1	BLAA_STRCT	P33651 streptomyc
853	4	40.0	308	1	KQGD_BACSU	P42235 bacillus su	926	4	40.0	326	1	TG12_STRCT	P34857 saccharomyc
854	4	40.0	308	1	URIC_CANLI	P34798 canavalia su	927	4	40.0	326	1	VP40_EBOZM	Q05128 ebola virus
855	4	40.0	308	1	URIC_PHAVU	P53763 phaseolus v	928	4	40.0	326	1	VS09_ROTHT	P16501 human rotav
856	4	40.0	308	1	YL58_CAEBL	P34438 caenorhabdi	929	4	40.0	326	1	VS09_ROTHT	P11856 human rotav
857	4	40.0	309	1	HEWZ_AQUAE	O67083 aquifex aeo	930	4	40.0	326	1	VS09_ROTHT	P08365 porcine rot
858	4	40.0	309	1	P2A_DROME	P23696 drosophila	931	4	40.0	326	1	WZB1_ECOLI	P76372 escherichia
859	4	40.0	309	1	TUS_ECOLI	P16525 escherichia	932	4	40.0	326	1	Y168_HUMAN	P50749 homo sapien
860	4	40.0	310	1	ATPG_SPIPL	P50006 spirulina p	933	4	40.0	327	1	CG65_HUMAN	O93375 homo sapien
861	4	40.0	311	1	PYDB_LACIL	P54322 lactococcus	934	4	40.0	327	1	FBR1_MOUSE	P33550 mus musculu
862	4	40.0	311	1	VORR_PYRHO	O58414 pyrococcus	935	4	40.0	327	1	GLSA_SYNY3	P73903 synecocyst
863	4	40.0	311	1	YBRU_BORBR	O06702 bordetella	936	4	40.0	327	1	OGA1_HUMAN	O95222 homo sapien
864	4	40.0	312	1	VP12_RDVA	O05054 rice dwarf	937	4	40.0	327	1	WZB8_SALTY	O04866 salmoneilla
865	4	40.0	312	1	VP12_RDVF	O05443 rice dwarf	938	4	40.0	327	1	Y745_HELPY	O94198 helicobacte
866	4	40.0	312	1	VP12_RDVO	O86480 rice dwarf	939	4	40.0	327	1	Y745_HELPY	O25441 helicobacte
867	4	40.0	312	1	YCXX_CHLMO	P09754 chlamydomon	940	4	40.0	327	1	Y1G3_YEAST	P40517 saccharomyc
868	4	40.0	313	1	CBBR_RHIME	P56885 rhizobium m	941	4	40.0	327	1	YN91_YEAST	P53749 saccharomyc
869	4	40.0	313	1	FWT_THEMEA	O94Y28 thermotoga	942	4	40.0	328	1	COBP_PSEEM	P33375 pseudomonas
870	4	40.0	314	1	HEWZ_CHLTP	O84492 chlamydia t	943	4	40.0	328	1	IBP2_HUMAN	P10065 homo sapien
871	4	40.0	314	1	MDH_RICPR	Q926F3 rickettsia	944	4	40.0	328	1	IRL1_HUMAN	O01638 homo sapien
872	4	40.0	314	1	SIX3_CHICK	O42406 gallus gall	945	4	40.0	328	1	RPOA_MYCBE	P47423 mycoplasma
873	4	40.0	315	1	GBF1_ARATH	P42774 arabidopsis	946	4	40.0	328	1	T2EB_YEAST	P36145 saccharomyc
874	4	40.0	315	1	HYFC_ECOLI	P77558 escherichia	947	4	40.0	328	1	YG19_YEAST	P53208 saccharomyc
875	4	40.0	315	1	MCH_METKA	P94954 methanopyru	948	4	40.0	329	1	COXA_STRCO	O66779 streptomyc
876	4	40.0	315	1	RPOA_CLOPE	O91BW9 clostridium	949	4	40.0	329	1	Y493_MYCTU	O11158 mycobacteri
877	4	40.0	315	1	RSEB_TAEIN	P44792 haemophilus	950	4	40.0	330	1	HEW2_RAT	O06214 rattus norv
878	4	40.0	316	1	MCH_ARCFU	O28344 archaeoglob	951	4	40.0	330	1	SWR_RICPR	O06214 rattus norv
879	4	40.0	316	1	YB33_YEAST	P38298 saccharomyc	952	4	40.0	330	1	Y677_METUA	O50890 methanococc
880	4	40.0	317	1	LDH_BACCA	P10655 bacillus ca	953	4	40.0	331	1	FHR4_HUMAN	O94496 homo sapien
881	4	40.0	317	1	RSP4_URECA	P38981 urechis cau	954	4	40.0	331	1	LEU1_THERM	O56216 thermus aqu
882	4	40.0	317	1	Y187_AQUAE	O67728 aquifex aeo	955	4	40.0	331	1	MA2_MESAU	P56670 mesocricetu
883	4	40.0	317	1	Y207_AQUAE	O66402 aquifex aeo	956	4	40.0	331	1	Y244_METUA	O60299 methanococc
884	4	40.0	318	1	NUIM_TAMTE	O78699 tamandua te	957	4	40.0	332	1	AXHA_ASPNG	P79019 aspergillus
885	4	40.0	318	1	PEX7_MOUSE	P97865 mus musculu	958	4	40.0	332	1	AXHA_ASPNG	P79021 aspergillus
886	4	40.0	319	1	HALL_RAT	P15978 rattus norv	959	4	40.0	332	1	LYNB_MYCLE	O95028 drosophila
887	4	40.0	319	1	LDH_BIFLO	P19869 bifidobacte	960	4	40.0	332	1	LYNB_MYCLE	O95028 drosophila
888	4	40.0	319	1	MCH_METTM	P51616 methanobact	961	4	40.0	332	1	STX3_HUMAN	O93781 mycobacteri
889	4	40.0	319	1	MOCB_SYNP7	Q56208 synecococc	962	4	40.0	332	1	SR4_PHYPO	O93343 homo sapien
890	4	40.0	319	1	STRN_STRGR	P29784 streptomyc	963	4	40.0	332	1	YCJW_ECOLI	P11113 physarum po
891	4	40.0	319	1	YPR5_CAEBL	P41883 caenorhabdi	964	4	40.0	333	1	MOSA_RHIME	P76165 escherichia
892	4	40.0	320	1	CHIX_PEA	P36007 pisum sativ	965	4	40.0	333	1	SIX3_MOUSE	O67607 rhizobium m
893	4	40.0	320	1	GSMB_SYNY3	P73493 synecocyst	966	4	40.0	334	1	CHMU_ARATH	O62233 mus musculu
894	4	40.0	320	1	KHSE_ZYMMO	O69015 zymomonas m	967	4	40.0	334	1	DPOB_HUMAN	P43738 arabidopsis
895	4	40.0	320	1	MCH_METTH	O26667 methanobact	968	4	40.0	334	1	DPOB_HUMAN	P06746 homo sapien
896	4	40.0	320	1	PR73_MMYVQ	P03320 mouse mamma	969	4	40.0	334	1	FLA3_PYRHO	O58284 pyrococcus
897	4	40.0	320	1	THRR_SYNP7	P27477 synecococc	970	4	40.0	334	1	MDHM_YEAST	P15505 saccharomyc
898	4	40.0	321	1	FBR1_HUMAN	P22087 homo sapien	971	4	40.0	334	1	Y111_HALNI	P23464 halobacteri
899	4	40.0	321	1	LIPA_ECOLI	P25845 escherichia	972	4	40.0	335	1	COLD_HUMAN	P15433 homo sapien
900	4	40.0	321	1	MCH_METBA	P94919 methanosarc	973	4	40.0	335	1	LYB1_MYCTU	O54358 mycobacteri
901	4	40.0	321	1	SAPB_ECOLI	Q47623 escherichia	974	4	40.0	335	1	TRPD_HELPY	O921J7 helicobacte
902	4	40.0	322	1	SAPB_SALTY	P36668 salmoneilla	975	4	40.0	335	1	TRPD_HELPY	P56737 helicobacte
903	4	40.0	322	1	ARCI_YEAST	P05089 homo sapien	976	4	40.0	335	1	TWR8_CAEBL	P34410 caenorhabdi
904	4	40.0	322	1	CCCL_YEAST	P47818 saccharomyc	977	4	40.0	336	1	DLDH_ACHLA	P33484 acholoplasm
905	4	40.0	322	1	HA10_MOUSE	P01898 mus musculu	978	4	40.0	336	1	Y625_METUA	O58042 methanococc
906	4	40.0	322	1	HCR_ECOLI	P75824 escherichia	979	4	40.0	337	1	LEPB_HAESO	P36685 haemophilus
907	4	40.0	323	1	SCI3_HUMAN	P55735 homo sapien	980	4	40.0	339	1	F16Q_HUMAN	O00757 homo sapien
908	4	40.0	323	1	FBR1_XENLA	P22332 xenopus lae	981	4	40.0	339	1	KDGT_ERMCH	P15701 erwinia chr
909	4	40.0	323	1	MCH_METEX	O85014 methyllobact	982	4	40.0	339	1	MURG_THEMEA	O94Y74 thermotoga

691	4	40.0	263	1	T2T8_THETH	P29748	thermus aqu	764	4	40.0	281	1	TRPA_METJA	Q60180	methanococ
692	4	40.0	263	1	TRPA_ANTSP	P31204	antilhamio	765	4	40.0	282	1	YC53_METJA	Q58650	methanococ
693	4	40.0	264	1	CXB1_XENIA	P08983	xenopus lae	766	4	40.0	283	1	Y43F_MYCPN	P75158	mycoplasma
694	4	40.0	264	1	DMPH_PSESP	P49156	pseudomonas	767	4	40.0	284	1	RP32_PROMI	P50509	proteus mir
695	4	40.0	264	1	IOD2_RANCA	P49896	rana catesb	768	4	40.0	284	1	VDLC_HELPY	O05730	helicobacte
696	4	40.0	264	1	PUL1_HUMAN	P17947	homo sapien	769	4	40.0	285	1	TONB_HELPY	O25999	helicobacte
697	4	40.0	264	1	RCIT_BPP4	P13059	bacterioph	770	4	40.0	286	1	MCMI_YEAST	P11746	saccharomy
698	4	40.0	264	1	RSP4_STRPU	P46771	strongyloce	771	4	40.0	287	1	MYOC_RAT	P20448	rattus norv
699	4	40.0	264	1	TRPA_SYNZ3	P77960	synecocyst	772	4	40.0	287	1	RL3_MYCPN	P75580	mycoplasma
700	4	40.0	264	1	TRPC_LACLA	O01999	lactococcus	773	4	40.0	287	1	Y280_MYCPN	P75580	mycoplasma
701	4	40.0	265	1	APAI_PIG	P18668	sus scrofa	774	4	40.0	287	1	Y1CC_HAEIN	P44726	haemophilus
702	4	40.0	265	1	DH47_ARATH	P31168	arabidopsis	775	4	40.0	288	1	EPMO_HUMAN	P32856	homo sapien
703	4	40.0	265	1	IOD2_HUMAN	O92813	homo sapien	776	4	40.0	288	1	FXE3_MOUSE	O99Y14	mus musculus
704	4	40.0	265	1	REFAP_ECOLI	P25741	escherichia	777	4	40.0	288	1	HMXX_CHICK	P50232	gallus gall
705	4	40.0	265	1	TRPA_METTM	P26920	methanobact	778	4	40.0	288	1	STIA_MOUSE	O35526	mus musculus
706	4	40.0	265	1	YMI2_PARTE	P15613	paramectum	779	4	40.0	288	1	WTL_ALUMI	P50992	alligator m
707	4	40.0	266	1	PUL1_MOUSE	P17433	mus musculus	780	4	40.0	288	1	Y940_MYCTU	P71359	mycobacteri
708	4	40.0	266	1	SURE_METJA	O57979	methanococ	781	4	40.0	289	1	ATPG_HAEIN	P43716	haemophilus
709	4	40.0	266	1	UXUR_HAEIN	P44487	haemophilus	782	4	40.0	289	1	DAPA_METJA	O57695	methanococ
710	4	40.0	267	1	STO3_CABEL	O20657	caenorhabdl	783	4	40.0	289	1	FRAP_ANASP	P46017	anabaena sp
711	4	40.0	267	1	TRPA_BACSU	P07601	bacillus su	784	4	40.0	289	1	SYLB_AQUAE	O67646	aquifex aeo
712	4	40.0	267	1	Y412_METJA	O57855	methanococ	785	4	40.0	290	1	ATPG_BACSU	P37810	bacillus su
713	4	40.0	268	1	ENGA_SALTY	O9XC18	salmonella	786	4	40.0	290	1	HXB8_HUMAN	P13378	homo sapien
714	4	40.0	268	1	KSGA_RICPR	O05952	r dimethyla	787	4	40.0	290	1	STX_APLCA	O16932	aplysia cal
715	4	40.0	269	1	NUDC_VIBCH	O9KX27	vibriol chol	788	4	40.0	290	1	TODE_PSEPU	P13453	pseudomonas
716	4	40.0	269	1	TRPA_BACST	P19867	bacillus st	789	4	40.0	291	1	YJBC_ECOLI	P32684	escherichia
717	4	40.0	270	1	HCE1_ORYLA	P31580	oryzias lat	790	4	40.0	291	1	C1LB_HAEIN	P44460	haemophilus
718	4	40.0	270	1	ILIA_MOUSE	P01582	mus musculus	791	4	40.0	291	1	GTH2_WHEAT	P30111	triticum ae
719	4	40.0	270	1	MOTB_BACSU	P28611	bacillus su	792	4	40.0	291	1	STXA_DROME	O24547	drosophila
720	4	40.0	270	1	PANB_THEMA	O9XZ31	thermotoga	793	4	40.0	292	1	COPE_CABEL	O62246	caenorhabdl
721	4	40.0	270	1	PYRD_THEMA	O9WY98	thermotoga	794	4	40.0	292	1	YVRE_BACSU	O34940	bacillus su
722	4	40.0	270	1	TRPA_METTH	O27697	methanobact	795	4	40.0	293	1	HVRI_MOUSE	P13297	mus musculus
723	4	40.0	270	1	YIDA_ECOLI	P09997	escherichia	796	4	40.0	293	1	META_CAMJE	O9P142	campylobact
724	4	40.0	271	1	RL7A_ANGA	O76732	anopheles g	797	4	40.0	293	1	Y326_MYCPN	P75312	mycoplasma
725	4	40.0	271	1	VDLC_HELPY	O9ZKX1	helicobacte	798	4	40.0	294	1	Z39F_HUMAN	O15777	homo sapien
726	4	40.0	271	1	Y228_TREPA	O83236	treponema p	799	4	40.0	294	1	DAPA_BUCAL	P57197	buchnera ap
727	4	40.0	271	1	YW34_MYCTU	O05879	mycobacteri	800	4	40.0	294	1	DAPA_THEMA	O9W2X1	thermotoga
728	4	40.0	272	1	VG67_HSYEB	P28964	equine herp	801	4	40.0	294	1	FLAI_PYRKO	P16233	haemonchus
729	4	40.0	272	1	YIS2_SHISO	P16940	shigella so	802	4	40.0	295	1	CAC3_HABCO	O58519	methanococ
730	4	40.0	273	1	GIAB_GIALA	P15518	giardia lam	803	4	40.0	295	1	DAPF_METJA	P15126	xenopus lae
731	4	40.0	273	1	TRPC_BRAJA	P94327	bradyrhizob	804	4	40.0	295	1	RL5B_XENLA	P15126	xenopus lae
732	4	40.0	274	1	PRCE_ARATH	O23717	arabidopsis	805	4	40.0	295	1	SUCD_PSEAE	O51567	pseudomonas
733	4	40.0	274	1	VGLE_HSV2	P13289	herpes simp	806	4	40.0	295	1	TRPC_SYNZ3	O55508	synecocyst
734	4	40.0	274	1	Y026_METJA	O60343	methanococ	807	4	40.0	295	1	YHOB_YEAST	P38805	saccharomyc
735	4	40.0	274	1	YBLC_SCYPO	O9URV0	schizosacch	808	4	40.0	295	1	110K_PLAKN	P13813	plasmotiomy
736	4	40.0	275	1	FLA5_PYRKO	O9V2X7	pyrococcus	809	4	40.0	296	1	CCO1_CABEL	P08124	caenorhabdl
737	4	40.0	275	1	VE39_NPYAC	P11042	autographa	810	4	40.0	296	1	METF_AQUAE	O67442	aquifex aeo
738	4	40.0	276	1	COAT_TNVA	P22959	tobacco nec	811	4	40.0	296	1	RL5_CHICK	P22451	gallus gall
739	4	40.0	276	1	LIPA_AQUAE	O67368	aquifex aeo	812	4	40.0	296	1	RL5_HUMAN	P09895	rattus norv
740	4	40.0	276	1	PLPB_PASHA	O08869	pasteurella	813	4	40.0	296	1	YH9F_HELPY	O92917	homo sapien
741	4	40.0	276	1	RSP4_CABEL	P46769	caenorhabdl	814	4	40.0	296	1	YH9F_HELPY	O92917	homo sapien
742	4	40.0	277	1	BHMT_PIG	O95332	sus scrofa	815	4	40.0	297	1	RL5_RAT	O02766	bos taurus
743	4	40.0	277	1	HXDB_NOTVI	P31263	notophtalm	816	4	40.0	297	1	HMX1_HUMAN	P28360	homo sapien
744	4	40.0	277	1	TRPA_HALVO	P18284	halobacteri	817	4	40.0	297	1	OCGR_RHIME	P72294	rhizobium m
745	4	40.0	278	1	Y31K_SSV1	P20203	sulfolobus	818	4	40.0	297	1	YNFL_ECOLI	P77559	escherichia
746	4	40.0	278	1	FIXR_BRAJA	P05406	bradyrhizob	819	4	40.0	298	1	DAPA_HAEIN	P43797	haemophilus
747	4	40.0	278	1	T2D7_DROME	O27272	drosophila	820	4	40.0	298	1	OCGR_AGRTU	O00679	agrobacteri
748	4	40.0	278	1	T2N2_NEIGO	P24617	neisseria g	821	4	40.0	298	1	Y33K_HUMAN	O04323	homo sapien
749	4	40.0	278	1	VGLL_HCMV1	O68657	human cytom	822	4	40.0	299	1	ATPG_ACEMO	P50005	acetobacter
750	4	40.0	278	1	VGLL_HCMV2	O68668	human cytom	823	4	40.0	299	1	Y175_HELPY	O92917	homo sapien
751	4	40.0	278	1	VGLL_HCMV3	O68668	human cytom	824	4	40.0	299	1	Y175_HELPY	O92917	homo sapien
752	4	40.0	278	1	VGLL_HCMV4	O68670	human cytom	825	4	40.0	300	1	YJH9_YEAST	P47032	saccharomyc
753	4	40.0	278	1	VGLL_HCMV5	O68671	human cytom	826	4	40.0	300	1	CEPM_BACHD	O92917	bacillus ha
754	4	40.0	278	1	VGLL_HCMV6	O68672	human cytom	827	4	40.0	301	1	CC21_TREBB	P38973	trypanosoma
755	4	40.0	278	1	VGLL_HCMV7	O68673	human cytom	828	4	40.0	301	1	FEAR_ECOLI	O47129	escherichia
756	4	40.0	278	1	VGLL_HCMV8	O68674	human cytom	829	4	40.0	301	1	NHAR_ECOLI	P10087	escherichia
757	4	40.0	278	1	VGLL_HCMV9	O68675	human cytom	830	4	40.0	301	1	YURD_CANLI	P34799	canavalia l
758	4	40.0	278	1	VGLL_HCMV10	O68676	human cytom	831	4	40.0	302	1	YMBI_BACSU	P39592	bacillus su
759	4	40.0	279	1	HCE2_ORYLA	P31581	oryzias lat	832	4	40.0	302	1	CASB_MACEU	P28550	macropus eu
760	4	40.0	281	1	DEGV_BACSU	P32436	bacillus su	833	4	40.0	302	1	C1LB_ECOLI	P77770	escherichia
761	4	40.0	281	1	HEMK_AQUAE	O66506	aquifex aeo	834	4	40.0	302	1	YEB6_MYCTU	P71766	mycobacteri
762	4	40.0	281	1	RM30_YEAST	P36528	saccharomyc	835	4	40.0	304	1	Y39A_HUMAN	O13442	homo sapien
763	4	40.0	281	1	T2MT_METTF	P29565	methanobact	836	4	40.0	304	1	HEV1_CANFA	O9T5Z2	canis fami

545	4	40.0	231	1	PCR_B_ARCFU	O29844 archaeoglob	618	4	40.0	249	1	PCNA_PYRAB	O9uyx8 pyrococcus
546	4	40.0	231	1	RSMB_MOUSE	P27048 mus musculu	619	4	40.0	249	1	PCNA_PYRFU	O73947 pyrococcus
547	4	40.0	232	1	CASB_PIG	P39037 sus scrofa	620	4	40.0	249	1	PCNA_PYRHO	O58369 pyrococcus
548	4	40.0	232	1	KUK_PIG	P00752 sus scrofa	621	4	40.0	249	1	REP2_ZYGBA	P13780 zygosacchar
549	4	40.0	233	1	BRAG_PSEAE	P21630 pseudomonas	622	4	40.0	249	1	REP2_ZYGBA	P13780 zygosacchar
550	4	40.0	233	1	PSMA_THEAC	P23156 thermoplasm	623	4	40.0	250	1	FRI_SOYBN	P43333 arabidopsis
551	4	40.0	233	1	RL1_THEMA	P29393 thermotoga h	624	4	40.0	250	1	TRPC_BACSU	P19976 glycine max
552	4	40.0	233	1	SERA_ENTHI	P21138 entamoeba h	625	4	40.0	250	1	Y627_PYRHO	P03964 bacillus su
553	4	40.0	234	1	GCH1_SYNY3	O55759 synechocyst	626	4	40.0	251	1	TPIS_LEIME	O58369 pyrococcus
554	4	40.0	234	1	RS3_ACTAC	P55827 actinobacill	627	4	40.0	251	1	TPIS_TYCR	P48498 leishmania
555	4	40.0	234	1	RS3_HAEIN	P43372 haemophilus	628	4	40.0	252	1	DUT_HUMAN	P32270 typanosoma
556	4	40.0	234	1	SODM_CANAL	O13401 candida alb	629	4	40.0	252	1	GSC_HUMAN	P33316 homo sapien
557	4	40.0	235	1	AURB_CHLNU	P21197 chloroflexu	630	4	40.0	253	1	EBSD_ENTFA	P56915 homo sapien
558	4	40.0	235	1	YHHQ_HAEIN	P44908 haemophilus	631	4	40.0	253	1	FRI_PEA	P36929 enterococcu
559	4	40.0	236	1	CD8A_RAT	P07725 rattus norv	632	4	40.0	253	1	HEMA_METHH	O19675 psium sativ
560	4	40.0	236	1	GAMT_HUMAN	O14353 homo sapien	633	4	40.0	253	1	KAD_CHLNU	O9pkx0 chlamydia m
561	4	40.0	236	1	GAMT_MOUSE	O39969 mus musculu	634	4	40.0	253	1	KDSB_HAEIN	P44490 haemophilus
562	4	40.0	236	1	YMI3_CAEEL	O21018 caenorhabdi	635	4	40.0	253	1	PMX2_HUMAN	O99811 homo sapien
563	4	40.0	237	1	COX2_TRIRO	O01556 trichophyto	636	4	40.0	253	1	RPSF_BACME	P51545 bacillus me
564	4	40.0	237	1	FLA1_PYRHO	O58281 pyrococcus	637	4	40.0	253	1	SQT_BACHD	O9K500 bacillus ha
565	4	40.0	237	1	GUN_ASPAC	P23669 aspergillus	638	4	40.0	253	1	SQT_BACSU	P37522 bacillus su
566	4	40.0	237	1	MTR4_METHH	O27227 methanobact	639	4	40.0	253	1	TPIS_COPDA	P37522 bacillus su
567	4	40.0	237	1	MTR4_METHH	P80184 methanobact	640	4	40.0	254	1	FRI_PHAVU	P18820 coplus japo
568	4	40.0	237	1	PSA5_SOYBN	O9m4t8 glycine max	641	4	40.0	254	1	TPIS_ARATH	P25699 phaseolus v
569	4	40.0	237	1	RL23_MYCPN	P75578 mycoplasma	642	4	40.0	254	1	TPIS_PETHY	P48491 arabidopsis
570	4	40.0	238	1	ERT3_HUMAN	P14138 homo sapien	643	4	40.0	254	1	YNO8_YEAST	P48495 petunia hyb
571	4	40.0	238	1	YR23_AGRVI	P70793 agrobacteri	644	4	40.0	254	1	FLA5_PYRHO	P53904 saccharomyc
572	4	40.0	239	1	RIM9_YEAST	O04734 saccharomyc	645	4	40.0	255	1	RL8A_YEAST	O58286 pyrococcus
573	4	40.0	239	1	UPPS_HAEIN	P4938 haemophilus	646	4	40.0	255	1	RL8B_YEAST	P17076 saccharomyc
574	4	40.0	239	1	WT1_SMIMA	P49953 smilthopsis	647	4	40.0	255	1	RNH2_BACSU	P29453 saccharomyc
575	4	40.0	240	1	B512_TRYCR	P18270 trypanosoma	648	4	40.0	255	1	RPSF_BACLI	O31744 bacillus su
576	4	40.0	240	1	FARR_ECOLI	P13669 escherichia	649	4	40.0	255	1	RPSF_BACLI	P6766 bacillus li
577	4	40.0	240	1	HOL_SYNY3	P72849 synechocyst	650	4	40.0	255	1	TONB_YEREN	P07860 bacillus su
578	4	40.0	240	1	RSMB_CHICK	O9pv94 gallus galli	651	4	40.0	255	1	TRPC_BACME	O05740 versinia en
579	4	40.0	240	1	RSMB_ERIEU	O9vu67 eritaceus e	652	4	40.0	255	1	YORE_PSECL	P70937 bacillus me
580	4	40.0	240	1	RSMB_HUMAN	P14678 homo sapien	653	4	40.0	256	1	GI02_RHOCA	O03003 pseudomonas
581	4	40.0	240	1	RSMB_MONDO	O9vu66 monodelphis	654	4	40.0	256	1	GSC_MOUSE	P69691 rhodobacter
582	4	40.0	240	1	RSMN_HUMAN	P14076 homo sapien	655	4	40.0	256	1	Y165_METJA	O02591 mus musculu
583	4	40.0	241	1	RASH_ECOLI	P01115 harvey murti	656	4	40.0	257	1	CG1C_ORYSA	O57650 methanococc
584	4	40.0	241	1	RASH_MSVHA	P57077 homo sapien	657	4	40.0	257	1	HS70_LUPOO	P33411 oryza sativ
585	4	40.0	242	1	FLGH_CAOCR	P22606 caulobacter	658	4	40.0	257	1	TRPC_MACFA	P16121 lupinus pol
586	4	40.0	242	1	CAVT_BRALA	P05548 brachiosteo	659	4	40.0	257	1	TRPC_AOUAE	O07276 macaca fasc
587	4	40.0	243	1	DENR_HUMAN	O43583 homo sapien	660	4	40.0	258	1	H1_YEAST	O67675 aquilex aeo
588	4	40.0	243	1	GSCB_XENLA	P29454 xenopus lae	661	4	40.0	258	1	HB21_SPAEH	P53551 saccharomyc
589	4	40.0	243	1	GSCB_XENLA	P29454 xenopus lae	662	4	40.0	258	1	KH31_PAPHA	P15465 spatax leuc
590	4	40.0	243	1	MURK_HAEIN	P51586 haemophilus	663	4	40.0	258	1	YH33_THEMA	O28773 papio hamad
591	4	40.0	243	1	TRUA_MYCPN	O050291 mycoplasma	664	4	40.0	259	1	GATR_ECOLI	O3x255 thermotoga
592	4	40.0	243	1	HIO_CHITH	O07134 chironomus	665	4	40.0	259	1	NAD2_METJA	P36930 escherichia
593	4	40.0	244	1	S40A_CHITE	O05018 chironomus	666	4	40.0	259	1	RMS2_ARATH	O68747 methanococc
594	4	40.0	244	1	YK91_MYCTU	O10700 mycobacteri	667	4	40.0	260	1	1433_LYCES	P42814 arabidopsis
595	4	40.0	244	1	YR01_CAEEL	O10014 caenorhabdi	668	4	40.0	260	1	143C_TORAC	P33209 lycopersico
596	4	40.0	245	1	KAD_CHLTR	O84130 chlamydia t	669	4	40.0	260	1	PSMA_THERK	P93343 nicotiana t
597	4	40.0	245	1	UNG_MYCSE	P43433 mycoplasma	670	4	40.0	260	1	SHP_MOUSE	O24733 thermococcu
598	4	40.0	245	1	YH77_ARCFU	O28497 archaeoglob	671	4	40.0	260	1	TRPC_LACCA	O62227 mus musculu
599	4	40.0	246	1	DOG1_YEAST	P38774 saccharomyc	672	4	40.0	261	1	DHG2_BACME	P17217 lactobacilli
600	4	40.0	246	1	KAD_ARATH	O82514 arabidopsis	673	4	40.0	261	1	KLK2_HUMAN	P39043 bacillus me
601	4	40.0	246	1	MCT2_SHEEP	P79204 ovis aries	674	4	40.0	261	1	KLK3_HUMAN	P20151 homo sapien
602	4	40.0	246	1	RI4_HALMA	O12733 haliaurcula	675	4	40.0	261	1	MOTB_BACSU	P07268 homo sapien
603	4	40.0	246	1	STX6_HUMAN	O95445 homo sapien	676	4	40.0	261	1	TNF5_BOVIN	P28612 bacillus su
604	4	40.0	247	1	GRAB_HUMAN	P10144 h granzyme	677	4	40.0	261	1	Y497_MYCLE	P51749 bos taurus
605	4	40.0	247	1	MCT1_HUMAN	P23946 homo sapien	678	4	40.0	262	1	YOB4_CAEEL	P54580 mycobacteri
606	4	40.0	247	1	MCT1_MACFA	P56433 macaca fasc	679	4	40.0	262	1	APAL_BRARE	O09292 caenorhabdi
607	4	40.0	247	1	MCT1_PAPHA	P52195 papio hamad	680	4	40.0	262	1	COR4_WHEAT	O42363 brachydanio
608	4	40.0	247	1	MCT2_MERUN	P50341 meriones un	681	4	40.0	262	1	IOD2_MOUSE	P46524 titlicum ae
609	4	40.0	247	1	MCT3_RAT	P21844 mus musculu	682	4	40.0	262	1	KLK1_HUMAN	O32119 mus musculu
610	4	40.0	247	1	PMX2_MOUSE	O06348 mus musculu	683	4	40.0	262	1	RS4Y_HUMAN	P70551 rattus norv
611	4	40.0	248	1	RS3_CAEEL	P48153 caenorhabdi	684	4	40.0	262	1	RS4_CHICK	P22090 homo sapien
612	4	40.0	248	1	TPIS_CIOAB	O52653 clostridium	685	4	40.0	262	1	RS4_CRICK	P47880 gallus galli
613	4	40.0	248	1	TPIS_LYCES	P33206 lycopersico	686	4	40.0	262	1	THIM_ECOLI	P47961 ciceretulus
614	4	40.0	249	1	HMX1_CHICK	P28361 gallus galli	687	4	40.0	262	1	MAZG_HAEIN	P12750 homo sapien
615	4	40.0	249	1	MCT1_CANFA	P21842 canis famli	688	4	40.0	263	1	SGS3_DROYA	P44733 haemophilus
616	4	40.0	249	1			689	4	40.0	263	1		P13728 drosophila
617	4	40.0	249	1			690	4	40.0	263	1		



399	4	40.0	191	1	MLE1_RABIT	P02602	oryctolagus	472	4	40.0	212	1	RAB2_RAT	P05712	rattus norv
400	4	40.0	191	1	PGHD_CANFA	Q9x665	canis fami1	473	4	40.0	212	1	RL14_HUMAN	P05014	homo sapien
401	4	40.0	192	1	GC42_HELPJ	Q9z142	helicobacte	474	4	40.0	212	1	RL1_AERPE	O9y9w6	aeropyrum p
402	4	40.0	192	1	GC42_HELPJ	008315	helicobacte	475	4	40.0	212	1	YTFB_ECOLI	P39310	escherichia
403	4	40.0	192	1	ZAK4_HUMAN	Q14206	homo sapien	476	4	40.0	213	1	RL14_RAT	P63507	rattus norv
404	4	40.0	193	1	H10_MOUSE	P10922	mus musculu	477	4	40.0	213	1	RL3_BACST	P28600	baecillus st
405	4	40.0	193	1	H10_MOUSE	P43278	rattus norv	478	4	40.0	213	1	UBC_ASPM2	P25619	african sw1
406	4	40.0	194	1	H5B_XENLA	P22845	xenopus lae	479	4	40.0	213	1	YBL5_YEAST	P38121	saccharomyc
407	4	40.0	194	1	PGHD_HORSE	O97921	equus cabal	480	4	40.0	214	1	FLA4_PYRHO	O58863	pyrococcus
408	4	40.0	194	1	RAS5_DICDI	P32254	dictyostell	481	4	40.0	214	1	NEUM_XENLA	P55860	xenopus lae
409	4	40.0	194	1	RS4_BOVIN	P79103	bos taurus	482	4	40.0	214	1	RL18_AERPE	O9y194	aeropyrum p
410	4	40.0	194	1	RS4_HORSE	P55832	equus cabal	483	4	40.0	214	1	RSMB_RAT	P17136	rattus norv
411	4	40.0	194	1	YOY8_CAEEL	P34665	caenorhabd1	484	4	40.0	214	1	SC14_SCHCO	P31575	schizophyll
412	4	40.0	195	1	HPEK_SYNY3	P72736	synecocyst	485	4	40.0	215	1	COAT_PMY	P16596	papaya mosa
413	4	40.0	195	1	TPIS_LACSA	P48493	lactuca sat	486	4	40.0	215	1	UBC_ASPB7	P27459	african sw1
414	4	40.0	195	1	YVZ3_CAEEL	P50437	caenorhabd1	487	4	40.0	215	1	Y256_HAEIN	P43973	haemophilus
415	4	40.0	196	1	AMEL_MOUSE	P45559	mus musculu	488	4	40.0	215	1	YK1_YEAST	P53144	saccharomyc
416	4	40.0	196	1	NTPE_ENTHR	P43436	enterococcu	489	4	40.0	216	1	KCY_CHLPN	O9z7y5	chlamydia p
417	4	40.0	197	1	HAMI_YEAST	P47119	saccharomyc	490	4	40.0	216	1	Y726_METJA	O581302	methanococc
418	4	40.0	197	1	TRPC_METJA	O57690	methanococc	491	4	40.0	217	1	FLA2_METJA	O58302	methanococc
419	4	40.0	198	1	PR73_MMTVC	P03319	mouse mamma	492	4	40.0	217	1	SOMA_CEREL	P56437	cervus elap
420	4	40.0	198	1	TNFA_MOUSE	P43488	mus musculu	493	4	40.0	218	1	CTA1_CAVPO	P81706	cavia porce
421	4	40.0	198	1	YV14_YEAST	P47089	saccharomyc	494	4	40.0	218	1	MEPB_BACSR	P16172	baecillus sp
422	4	40.0	199	1	NHAA_RHOSO	O53118	rhodococcus	495	4	40.0	218	1	MOG1_YEAST	P47123	saccharomyc
423	4	40.0	200	1	SSUE_PSEPU	O85762	pseudomonas	496	4	40.0	219	1	FLA4_PYRAB	O9y413	pyrococcus
424	4	40.0	201	1	COAT_LVX	P27334	lilly vltus	497	4	40.0	219	1	FLA4_PYRKO	O9v7w8	pyrococcus
425	4	40.0	201	1	LEUD_ECOLI	P30126	escherichia	498	4	40.0	219	1	H1B_XENLA	O28272	archaeoglob
426	4	40.0	201	1	PYRE_HELPJ	O9z1x0	helicobacte	499	4	40.0	219	1	RI07_ARCFU	O28869	archaeoglob
427	4	40.0	201	1	PYRE_HELPJ	P56162	helicobacte	500	4	40.0	222	1	AMTB_SECCO	P30771	secale cere
428	4	40.0	201	1	YPT1_PVTIN	O01690	phytophthor	501	4	40.0	222	1	KCY2_HAEIN	O43963	haemophilus
429	4	40.0	202	1	AMEL_MONDO	Q28462	monodelphis	502	4	40.0	223	1	RNS2_SOLTU	O01796	solanum tub
430	4	40.0	202	1	COAT_ELV	P35927	erysimum la	503	4	40.0	224	1	YDHC_BACSU	O05494	baecillus su
431	4	40.0	202	1	FTHC_HUMAN	P49914	homo sapien	504	4	40.0	224	1	ASR2_CRIGR	O60436	bractetolopa
432	4	40.0	202	1	HR2A_TRIFL	P14530	trifleresuru	505	4	40.0	225	1	GP30_BRSP1	O38423	bacterioph
433	4	40.0	202	1	SDC4_RAT	P34901	rattus norv	506	4	40.0	225	1	RS3_LEPIN	O9x330	leptospira
434	4	40.0	204	1	BTB3_MOUSE	O64152	mus musculu	507	4	40.0	225	1	YB67_AERPE	O9y430	aeropyrum p
435	4	40.0	204	1	CLH5_CLAHE	P42059	cladospoirlu	508	4	40.0	226	1	14P_HUMAN	O14530	homo sapien
436	4	40.0	205	1	DUT_RAT	P70583	rattus norv	509	4	40.0	226	1	ATP6_HELPJ	O9z115	helicobacte
437	4	40.0	205	1	HYFA_ECOLI	P23481	escherichia	510	4	40.0	226	1	ATP6_HELPJ	P56605	helicobacte
438	4	40.0	205	1	KAD_MICLU	P33107	micrococcus	511	4	40.0	226	1	ATP6_HELPJ	O01704	brachydanio
439	4	40.0	205	1	RG54_HUMAN	P49798	homo sapien	512	4	40.0	226	1	HMMD_BRARE	O96184	mycobacteri
440	4	40.0	205	1	RG54_MOUSE	O08899	mus musculu	513	4	40.0	226	1	LP07_MYCTU	O47670	chrysocyon
441	4	40.0	205	1	RG54_RAT	P49799	rattus norv	514	4	40.0	226	1	OG61_AOUAE	O47668	cuon alpinu
442	4	40.0	205	1	YN50_YEAST	P42844	saccharomyc	515	4	40.0	227	1	COX2_CANAD	O47672	dusicyon th
443	4	40.0	206	1	Y086_CAEEL	P34622	caenorhabd1	516	4	40.0	227	1	COX2_CANAU	P48890	fells silve
444	4	40.0	207	1	FLA2_PYRHO	O58283	pyrococcus	517	4	40.0	227	1	COX2_CANME	O47671	canis aureu
445	4	40.0	207	1	IL6_MARMO	O35736	marrocta mon	518	4	40.0	227	1	COX2_CANST	P98031	canis simen
446	4	40.0	207	1	LEXA_AERHY	O44069	aeromonas h	519	4	40.0	227	1	COX2_CHRR	O47670	chrysocyon
447	4	40.0	207	1	PLCR_PSEAE	P40695	pseudomonas	520	4	40.0	227	1	COX2_CUOAL	O47668	cuon alpinu
448	4	40.0	208	1	CSF3_MOUSE	P09920	mus musculu	521	4	40.0	227	1	COX2_DUSTH	O47672	dusicyon th
449	4	40.0	208	1	GC43_HUMAN	O95843	homo sapien	522	4	40.0	227	1	COX2_FELCA	P48890	fells silve
450	4	40.0	208	1	NOIY_RHISN	P55716	rhicobium s	523	4	40.0	227	1	COX2_HAAGR	P38596	haichoerus
451	4	40.0	208	1	RS3A_SULSO	O9uxd4	sulfolobus	524	4	40.0	227	1	COX2_LVCP1	O47674	lycaon pict
452	4	40.0	208	1	SAOX_RHOCA	O52671	rhodobacter	525	4	40.0	227	1	COX2_MACRO	P92662	macropus ro
453	4	40.0	209	1	SPC3_STRPU	P16537	strongyloce	526	4	40.0	227	1	COX2_NYCPR	O47675	nyctereutes
454	4	40.0	209	1	RPOC_BUCAP	P42920	baecillus su	527	4	40.0	227	1	COX2_PHOVI	O00528	phoca vitul
455	4	40.0	209	1	YC02_KLEPN	P41185	buchnera ap	528	4	40.0	227	1	COX2_PSECP	O47676	pseudolopex
456	4	40.0	209	1	YC02_KLEPN	O48448	klebsiella	529	4	40.0	227	1	COX2_PSEEG	O47677	pseudolopex
457	4	40.0	209	1	YDGI_BACSU	P96707	baecillus su	530	4	40.0	227	1	COX2_PSESU	O47678	pseudolopex
458	4	40.0	209	1	YLZ3_CAEEL	P34416	caenorhabd1	531	4	40.0	227	1	COX2_SPEVA	O47679	specthos ve
459	4	40.0	210	1	RS7_HALHA	P15763	halobacteri	532	4	40.0	227	1	COX2_VOLMA	O47680	vulpes macr
460	4	40.0	210	1	TIPI_YEAST	P27654	saccharomyc	533	4	40.0	227	1	COX2_VULPE	O47681	vulpes vulp
461	4	40.0	211	1	Y121_HUMAN	O14135	homo sapien	534	4	40.0	227	1	COX2_VULZE	O47682	vulpes zerd
462	4	40.0	211	1	DEDD_ECOLI	P09549	escherichia	535	4	40.0	227	1	FLPA_PYRHO	O57861	pyrococcus
463	4	40.0	211	1	GLP1_STNAL	P45854	sinapis alb	536	4	40.0	227	1	KCY_ECOLI	P23863	escherichia
464	4	40.0	211	1	IL6_MOUSE	P08505	mus musculu	537	4	40.0	227	1	SYL_PSESY	P52833	pseudomonas
465	4	40.0	211	1	SERB_METJA	O58989	methanococc	538	4	40.0	227	1	T2H2_HAEHA	P00643	haemophilus
466	4	40.0	211	1	TX10_HUMAN	O75333	homo sapien	539	4	40.0	228	1	CAPB_STRAU	P39851	staphylococ
467	4	40.0	212	1	IFEA_CAEEL	O22888	caenorhabd1	540	4	40.0	229	1	Y4M6_RHISN	P55561	rhicobium s
468	4	40.0	212	1	RAB2_HUMAN	P08886	homo sapien	541	4	40.0	230	1	RT07_MARPO	P26667	marcho
469	4	40.0	212	1	RAB2_LYMS	O05975	lymaea sta	542	4	40.0	230	1	YP2A_STRAU	P03866	staphylococ
470	4	40.0	212	1	RAB2_MOUSE	P53994	mus musculu	543	4	40.0	231	1	FLGH_PSEPU	O52081	pseudomonas
471	4	40.0	212	1	RAB2_RABIT	O01971	oryctolagus	544	4	40.0	231	1	KCY1_HAEIN	P43892	haemophilus

253	4	40.0	142	1	HBAA_ANGAN	P80945	angullia an	326	4	40.0	163	1	ILVH_SALTY	P21622	salmonella
254	4	40.0	142	1	IL3_CALJA	Q28334	callitrix	327	4	40.0	163	1	SP17_PAPHA	G95230	papio hamad
255	4	40.0	142	1	IL3_SAGOE	P51445	sagunus oe	328	4	40.0	164	1	FLAV_HELPY	G94553	helicobacte
256	4	40.0	143	1	PER_DROPT	Q25206	drosophila	329	4	40.0	164	1	FLAV_HELPY	Q25276	helicobacte
257	4	40.0	143	1	YS85_CAEEL	Q09623	caenorhabdi	330	4	40.0	164	1	IM17_SCHPO	P81170	schizosacch
258	4	40.0	144	1	CSF2_CEREL	P51748	cervus elap	331	4	40.0	164	1	OV17_ONCVO	P36991	onchocerca
259	4	40.0	144	1	CSF2_SHEEP	P28773	ovis aries	332	4	40.0	164	1	Y0J3_CAEEL	P34626	caenorhabdi
260	4	40.0	144	1	HEX9_ADE12	P03284	human adeno	333	4	40.0	165	1	DEST_CHICK	P18359	gallus gall
261	4	40.0	144	1	RS14_TRYBB	P19800	trypanosoma	334	4	40.0	165	1	H1_TEMPY	P12305	tetrahymena
262	4	40.0	145	1	DH11_GOSHI	P09442	gospypium h	335	4	40.0	165	1	RS13_SULSO	P55968	sulfolobus
263	4	40.0	145	1	R19E_METHH	O27653	methanobact	336	4	40.0	165	1	RS41_MACBU	P79188	macaca tusc
264	4	40.0	145	1	YNR9_YEAST	P53880	saccharomyc	337	4	40.0	166	1	COF1_HUMAN	P23528	homo sapien
265	4	40.0	146	1	FLI3_BACSU	P20487	bacillus su	338	4	40.0	166	1	COF1_MOUSE	P18760	mus muscullu
266	4	40.0	147	1	PHEB_PAREP	P80721	paramphisto	339	4	40.0	166	1	COF1_PIG	P10668	sus scrofa
267	4	40.0	147	1	PHEB_BACSU	P21204	bacillus su	340	4	40.0	166	1	COF1_RAT	P45592	rattus norv
268	4	40.0	148	1	C552_THERH	P04164	thermus aqu	341	4	40.0	166	1	TPX_BACSU	P80864	bacillus su
269	4	40.0	148	1	CNRH_ALCEU	P37975	alcaligenes	342	4	40.0	167	1	QCRH_BACSU	P46911	bacillus su
270	4	40.0	148	1	YXIE_BACSU	P42297	bacillus su	343	4	40.0	167	1	YK63_YEAST	P35723	saccharomyc
271	4	40.0	149	1	LRRP_BACSU	P96653	bacillus su	344	4	40.0	169	1	GP38_CANFA	G95151	methanococ
272	4	40.0	149	1	MLE3_MOUSE	P05978	mus muscullu	345	4	40.0	169	1	Y1ST_BACSU	O07939	bacillus su
273	4	40.0	149	1	MLE3_RABIT	P02603	oryctolagus	346	4	40.0	170	1	FMF7_ECOLI	P25394	escherichia
274	4	40.0	149	1	MLE3_RAT	P02601	rattus norv	347	4	40.0	170	1	HEMX_PROXI	O51887	proteus mir
275	4	40.0	149	1	R19_HELPY	Q921W3	helicobacte	348	4	40.0	171	1	ATPF_HELPY	G94577	helicobacte
276	4	40.0	149	1	SP17_MOUSE	Q62252	mus muscullu	349	4	40.0	171	1	ATPF_HELPY	P36086	helicobacte
277	4	40.0	150	1	HS11_ORYSA	P27777	oryza sativ	350	4	40.0	171	1	IF3_LISMO	O53084	listeria mo
278	4	40.0	150	1	R19_HELPY	P56035	helicobacte	351	4	40.0	171	1	NUPM_BOVIN	P42029	bos taurus
279	4	40.0	150	1	RS13_METJA	P54019	methanococ	352	4	40.0	172	1	HS20_NIPBR	O07160	nippostrong
280	4	40.0	151	1	HS11_WHEAT	P12810	tritlicum ae	353	4	40.0	172	1	RUB2_PSEOL	P00272	pseudomonas
281	4	40.0	151	1	SOD1_LYCES	P14830	lycopersico	354	4	40.0	173	1	Y150_METJA	G95618	methanococ
282	4	40.0	151	1	SP17_MACFA	O19021	macaca fasc	355	4	40.0	174	1	HSGB_BUCAI	P57619	buchnera ap
283	4	40.0	152	1	CORA_RAT	O63532	rattus norv	356	4	40.0	174	1	NEUW_RAT	P12760	rattus norv
284	4	40.0	152	1	PYRI_PYRHO	O58452	pyrococcus	357	4	40.0	174	1	RS11_WHEAT	P00871	tritlicum ae
285	4	40.0	152	1	RL9_SYNY3	P42352	synechocyst	358	4	40.0	174	1	RBS_HORVU	O40004	hordeum vul
286	4	40.0	153	1	H2B1_CHLRE	P50565	chlamydomon	359	4	40.0	174	1	VC17_BPPEA	P08389	bacterioph
287	4	40.0	153	1	HS13_SOYBN	P04793	glycine max	360	4	40.0	174	1	YGT3_YEAST	P53102	saccharomyc
288	4	40.0	153	1	MYG_CARCR	P56208	caletta car	361	4	40.0	175	1	RBS2_WHEAT	P26667	tritlicum ae
289	4	40.0	153	1	MYG_CHEMY	P02202	chelonla my	362	4	40.0	175	1	RBS_AEGSO	G938791	aequilops sq
290	4	40.0	153	1	MYG_GRAGE	P02201	grapiemys g	363	4	40.0	175	1	YGT2_YEAST	P53054	saccharomyc
291	4	40.0	153	1	VC17_BPPE5	P15850	bacterioph	364	4	40.0	176	1	IPYR_AERPE	G93945	aeropyrum p
292	4	40.0	153	1	YPR4_CHRYI	P45373	chromatium	365	4	40.0	176	1	ITRT_ACACO	P24924	acacia conf
293	4	40.0	154	1	HS11_LYCES	P30221	lycopersico	366	4	40.0	176	1	VEGP_HUMAN	P31023	homo sapien
294	4	40.0	154	1	HS14_SOYBN	P04794	glycine max	367	4	40.0	176	1	Y017_MYCCE	P47163	mycoplasma
295	4	40.0	154	1	HS15_SOYBN	P04795	glycine max	368	4	40.0	177	1	APT_MYCPN	P75348	mycoplasma
296	4	40.0	155	1	BCCP_HAEIN	P43874	haemophilus	369	4	40.0	177	1	CYSC_SYNY3	P72940	synechocyst
297	4	40.0	155	1	PRL_PHAVU	P25985	phaseolus v	370	4	40.0	179	1	14P_BOVIN	O18883	bos taurus
298	4	40.0	155	1	PR2_PHAVU	P25986	phaseolus v	371	4	40.0	179	1	RK27_TOBRC	P30155	nicotiana t
299	4	40.0	155	1	RNH_SALTY	P23329	salmonella	372	4	40.0	180	1	KCY_ARCFU	O23023	archaeoglob
300	4	40.0	156	1	CUS5_ARADI	P80518	araneus dia	373	4	40.0	180	1	YC47_METJA	G95043	schistosoma
301	4	40.0	156	1	MAFK_HUMAN	O60675	homo sapien	374	4	40.0	181	1	NADD_CAMJE	Q58644	methanococ
302	4	40.0	156	1	Y723_METJA	G56133	methanococ	375	4	40.0	181	1	Y856_METJA	O58193	campylobact
303	4	40.0	157	1	HS2C_CHLRE	P18111	chlamydomon	376	4	40.0	183	1	RL17_PODCA	O58266	methanococ
304	4	40.0	157	1	PRTC_CANFA	Q28278	canis famill	377	4	40.0	183	1	STL_SULSO	P37380	podocoryne
305	4	40.0	157	1	VE6_HPV36	P50810	human papil	378	4	40.0	184	1	CAS3_MOUSE	O33768	sulfolobus
306	4	40.0	158	1	HS12_MEDSA	P27880	medicago sa	379	4	40.0	184	1	RL13_SCHMA	Q02862	mus muscullu
307	4	40.0	158	1	ILVH_BUCAP	O85294	buchnera ap	380	4	40.0	184	1	SM21_SCHMA	G95043	schistosoma
308	4	40.0	158	1	SAM2_SOYBN	P26987	glycine max	381	4	40.0	185	1	DE1_MOUSE	O60870	schistosoma
309	4	40.0	159	1	SH3B_DROME	O9nf05	drosophila	382	4	40.0	185	1	ERPL_CHLTR	O60870	mus muscullu
310	4	40.0	159	1	CUS7_ARADI	P80519	araneus dia	383	4	40.0	185	1	SC6_SCHCO	O44124	chlamydia t
311	4	40.0	159	1	DUT_OREN2	P14597	ori virus (	384	4	40.0	186	1	ARDH_LEIDO	O74300	schizophyll
312	4	40.0	159	1	Y268_BORBU	O44756	borrella bu	385	4	40.0	186	1	IF3_BORBU	Q05885	leishmania
313	4	40.0	160	1	PETD_ODOSI	P49489	odontella s	386	4	40.0	186	1	Y730_METJA	O51208	borrella bu
314	4	40.0	160	1	RL10_METHH	O27191	methanobact	387	4	40.0	187	1	BNWZ_PSEPU	O58140	methanococ
315	4	40.0	160	1	RL21_CAEEL	P34334	caenorhabdi	388	4	40.0	187	1	DCMB_PSEPU	P08085	pseudomonas
316	4	40.0	161	1	HS16_SOYBN	P05478	glycine max	389	4	40.0	187	1	LEPU_BACSU	P26693	methanotri
317	4	40.0	161	1	ID4_MOUSE	P41139	mus muscullu	390	4	40.0	187	1	MAUE_METEX	P42959	bacillus su
318	4	40.0	161	1	LE23_ARCFU	O29626	archaeoglob	391	4	40.0	187	1	MLEI_MOUSE	O49123	methylobact
319	4	40.0	161	1	VG41_HSVSA	O01027	herpesvirus	392	4	40.0	187	1	TOD2_PSEPU	P05977	mus muscullu
320	4	40.0	162	1	MAFG_HUMAN	O15525	homo sapien	393	4	40.0	187	1	Y160_THEMA	P13451	pseudomona
321	4	40.0	162	1	MAFG_MOUSE	O54790	mus muscullu	394	4	40.0	188	1	MLEI_RAT	G93244	thermotoga
322	4	40.0	163	1	HL_TETTH	P10136	tetrahymena	395	4	40.0	189	1	TNAE_HUMAN	P02600	rattus norv
323	4	40.0	163	1	HMCS_DICDI	P54872	dicyostell	396	4	40.0	189	1	PAAD_RICPR	P36609	caenorhabdi
324	4	40.0	163	1	ILVH_ECOLI	P00894	escherichia	397	4	40.0	189	1	PAAD_RICPR	O92609	rickettsia
325	4	40.0	163	1	ILVH_HAEIN	P45260	haemophilus	398	4	40.0	190	1	YBAY_ECOLI	P77717	escherichia

107	5	50.0	1265	1	PIG2_RAT	P24135	rattus norv	180	4	40.0	105	1	YMO3_CAEEL	P34494	caenorhabdl
108	5	50.0	1290	1	XCPC_XENLA	P50532	xenopus lae	181	4	40.0	106	1	CYC3_DESSA	P00135	desulfotivbr
109	5	50.0	1337	1	P152_YEAST	P39685	ratuococmyc	182	4	40.0	106	1	KACH_RABIT	P01393	oryctolagus
110	5	50.0	1472	1	A2MG_RAT	P06238	rattus norv	183	4	40.0	106	1	TBCH_YEAST	P48606	saccharomyc
111	5	50.0	1500	1	SSP5_STRGN	P16952	streptococ	184	4	40.0	107	1	U139_ARATH	Q96888	arabidopsis
112	5	50.0	1627	1	ADP1_MPCPN	P11311	mycoplasma	185	4	40.0	109	1	PRVA_MOUSE	P32848	mus musculu
113	5	50.0	1912	1	VIT1_CHICK	P87498	gallus galli	186	4	40.0	109	1	RBS_PROHO	P27569	prochloroth
114	5	50.0	1928	1	MYST_YEAST	P08964	saccharomyc	187	4	40.0	110	1	RPOW_THECE	Q56253	thermococc
115	5	50.0	2208	1	POLN_MANCV	Q69014	manchester	188	4	40.0	110	1	RS6_AOUAE	Q56474	aquifex aeo
116	5	50.0	2476	1	ZAN_PIG	Q28983	sus scrofa	189	4	40.0	110	1	Y16F_BPT4	P22916	bacterioph
117	5	50.0	3133	1	HMCT_BOMMO	P98092	bombyx mori	190	4	40.0	111	1	CCR4_STNY3	P73407	synchocyst
118	4	40.0	21	1	CXGT_CONTU	P17684	conus tulip	191	4	40.0	112	1	CL34_DROMI	Q01174	drosophila
119	4	40.0	32	1	ATP7_SPTOL	P80088	spinacia ol	192	4	40.0	113	1	HCCL_CRYCO	Q01239	cryptocodi
120	4	40.0	42	1	TFXA_RHLIT	P42723	rhizobium l	193	4	40.0	113	1	RBS3_MHEAT	P07398	triticum ae
121	4	40.0	44	1	Y5K3_VACCV	P18379	vaccinia vl	194	4	40.0	113	1	RLAI_DICDI	P22684	dictyostell
122	4	40.0	49	1	R332_BACLI	P35870	bacillus l1	195	4	40.0	113	1	Y12K_BPPH	P23788	bacterioph
123	4	40.0	49	1	R332_BACSU	Q06798	bacillus su	196	4	40.0	114	1	CU05_ROMAM	P81389	homoar ame
124	4	40.0	50	1	ITR2_BOVIN	P56651	bos taurus	197	4	40.0	114	1	ET3_RABIT	P19998	oryctolagus
125	4	40.0	53	1	ADO_RABIT	P80456	oryctolagus	198	4	40.0	115	1	GUAN_RAT	P28902	rattus norv
126	4	40.0	56	1	Y546_METJA	Q57966	methanococ	199	4	40.0	116	1	GCAD_BACCL	P42817	bacillus ca
127	4	40.0	59	1	TAT2_AOUAE	Q66477	aquifex aeo	200	4	40.0	116	1	GUAN_MOUSE	P33680	mus musculu
128	4	40.0	62	1	UCRX_BOVIN	P00130	bos taurus	201	4	40.0	116	1	RM14_ACAP0	P34829	acanthamoeb
129	4	40.0	65	1	YHSB_CIOAB	Q04333	clostridium	202	4	40.0	116	1	SP21_BACCO	P70877	bacillus co
130	4	40.0	65	1	YNOL_RHLER	P33214	rhizobium f	203	4	40.0	117	1	HV1A_HUMAN	P01742	homo sapien
131	4	40.0	66	1	YGLR_STRCO	Q05954	streptomyce	204	4	40.0	117	1	HV1B_HUMAN	P01743	homo sapien
132	4	40.0	66	1	YYCD_BACSU	P37480	bacillus su	205	4	40.0	117	1	HLVG_HUMAN	P23083	homo sapien
133	4	40.0	67	1	HJAI_METJA	Q57632	methanococ	206	4	40.0	117	1	NLTS_ORYZA	Q65091	oryza sativ
134	4	40.0	67	1	HJAI_METJA	Q58342	methanococ	207	4	40.0	119	1	NLRP_SALTI	Q56131	salmonella
135	4	40.0	67	1	HJAI_METJA	Q58655	methanococ	208	4	40.0	119	1	RL2A_LEPTN	Q94235	leptospira
136	4	40.0	67	1	HJAI_METJA	Q60264	methanococ	209	4	40.0	120	1	CU19_ARADI	Q90155	araneus dla
137	4	40.0	68	1	TRPC_BACPU	Q9WVW8	streptococ	210	4	40.0	121	1	YHIT_AOUAE	Q66536	aquifex aeo
138	4	40.0	68	1	RL29_STRPN	P18268	bacillus pu	211	4	40.0	121	1	RBS1_SPTOL	P00870	spinacia ol
139	4	40.0	72	1	VG35_BPMU5	Q05245	mycobacteri	212	4	40.0	123	1	Y949_METJA	Q59455	methanococ
140	4	40.0	73	1	Y556_ARCFU	Q29665	archaeoglob	213	4	40.0	124	1	GC5H_THEMA	Q9W459	thermotoga
141	4	40.0	73	1	YB52_METJA	Q58552	methanococ	214	4	40.0	124	1	HVID_HUMAN	P01760	homo sapien
142	4	40.0	75	1	YCCJ_ECOLI	P46131	escherichia	215	4	40.0	124	1	IBB_HORVU	P01761	homo sapien
143	4	40.0	77	1	KEA2_ECOLI	P31964	escherichia	216	4	40.0	124	1	HV1E_HUMAN	P12940	hordium vul
144	4	40.0	78	1	DSVD_DESVH	Q46582	desulfotivbr	217	4	40.0	124	1	SGP2_CHREVI	Q52119	chromatium
145	4	40.0	78	1	Y113_RICPR	Q9Z635	ricketsia	218	4	40.0	125	1	FAB2_AMBME	P81400	ambystoma m
146	4	40.0	82	1	RS16_SALTY	P36242	salmonella	219	4	40.0	125	1	FABL_CHICK	P80262	gallus gall
147	4	40.0	82	1	TATA_VIRCH	P57051	vibrio chol	220	4	40.0	125	1	FLIO_SALTY	Q54699	salmonella
148	4	40.0	83	1	GVG1_HALNI	P24371	halobacteri	221	4	40.0	125	1	HV1F_HUMAN	P06326	homo sapien
149	4	40.0	88	1	Y67_BPT7	P03801	bacterioph	222	4	40.0	125	1	SDIS_COMTE	P00516	comamonas t
150	4	40.0	89	1	ACYF_ARCFU	Q29440	archaeoglob	223	4	40.0	126	1	CU24_ARADI	P80517	araneus dla
151	4	40.0	90	1	LPPL_HUMAN	Q95968	homo sapien	224	4	40.0	127	1	CU26_ARADI	P44455	haemophilus
152	4	40.0	90	1	RPOL_SULAC	P46217	sulfobolus	225	4	40.0	127	1	YFID_HAEIN	Q83268	treponema p
153	4	40.0	90	1	Y032_HELPU	Q9Z932	helicobacte	226	4	40.0	129	1	RL7_TREPA	P46767	acanthamoeb
154	4	40.0	92	1	RL5_MOUSE	P47962	mus musculu	227	4	40.0	129	1	RM14_ACACA	P50618	bacillus su
155	4	40.0	93	1	IM09_ARATH	Q9X9X9	arabidopsis	228	4	40.0	130	1	NRDI_BACSU	Q91112	mus musculu
156	4	40.0	93	1	VOIC_CABEL	Q09283	caenorhabdl	229	4	40.0	130	1	SY28_MOUSE	Q57835	methanococ
157	4	40.0	94	1	NLPD_SALDU	P39700	salmonella	230	4	40.0	130	1	Y390_METJA	Q58081	methanococ
158	4	40.0	94	1	YNGC_CIOPE	P26834	clostridium	231	4	40.0	130	1	Y614_METJA	P71993	mycobacteri
159	4	40.0	96	1	SASG_BACME	P02961	bacillus me	232	4	40.0	130	1	Y056_MYCTU	P48312	human adeno
160	4	40.0	96	1	YCA_TYLCU	P38612	tomato yell	233	4	40.0	132	1	HEX9_ADE40	Q58265	methanococ
161	4	40.0	98	1	CYTB_MOOSE	Q62426	mus musculu	234	4	40.0	132	1	IF5A_METJA	Q05221	mycobacteri
162	4	40.0	98	1	NLPD_MOOSE	P40827	salmonella	235	4	40.0	132	1	VG24_BPMU5	C09191	schizosacch
163	4	40.0	98	1	YCA_TYLCM	P27272	tomato yell	236	4	40.0	133	1	CDD_SCHPO	P32599	human adeno
164	4	40.0	99	1	GP45_BPSPI	Q48339	bacterioph	237	4	40.0	133	1	HEX9_ADE41	P55030	fells alive
165	4	40.0	99	1	IF1A_METH	Q27085	methanobact	238	4	40.0	133	1	IL4_FELCA	Q81924	heliolantus
166	4	40.0	99	1	RL12_METVA	P10623	methanococ	239	4	40.0	133	1	PROF_HELAN	P40294	entamoeba 1
167	4	40.0	99	1	Y13J_BPT4	P39503	bacterioph	240	4	40.0	134	1	H2B_ENTY	Q99428	homo sapien
168	4	40.0	99	1	YORR_PYRMO	P20297	pyrococcus	241	4	40.0	134	1	VPD4_HUMAN	Q64214	homo sapien
169	4	40.0	100	1	REGN_BPP22	P04891	bacterioph	242	4	40.0	134	1	VG24_BPPM2	P31776	zea mays (m
170	4	40.0	101	1	RL5_RABIT	P19949	oryctolagus	243	4	40.0	135	1	CYT1_MAIZE	P38322	saccharomyc
171	4	40.0	102	1	SGP3_CHREVI	Q52055	chromatium	244	4	40.0	136	1	YB76_YEAST	Q63855	canis famill
172	4	40.0	102	1	VE7_PAPVE	P11332	european el	245	4	40.0	137	1	COX2_CANFA	P56265	klebsiella
173	4	40.0	103	1	ACBP_ANAPL	P45882	anas platyr	246	4	40.0	137	1	NIFB_KLEOX	Q9V091	pyrococcus
174	4	40.0	103	1	LAC_CHICK	P20763	gallus galli	247	4	40.0	138	1	GC5H_PYRAB	Q08052	pleurochrys
175	4	40.0	104	1	GAST_BOVIN	P01352	bos taurus	248	4	40.0	139	1	NEDR_MICVI	Q08052	pleurochrys
176	4	40.0	104	1	GAST_SHEEP	Q02666	ovis aries	249	4	40.0	139	1	RS6_PLECA	Q51142	borrelia bu
177	4	40.0	104	1	VR7H_ECOLI	P21332	escherichia	250	4	40.0	139	1	RS6_BORBU	Q25775	helicobacte
178	4	40.0	105	1	LAC_HUMAN	P01842	homo sapien	251	4	40.0	140	1	YB60_HELPU	Q04625	clostridium
179	4	40.0	105	1	NDC1_BRAVA	P26024	bradyrhizob	252	4	40.0	141	1	YFE3_CLOPA		

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:43:05 ; Search time 20.38 Seconds  
(without alignments)  
16.808 Million cell updates/sec

Title: US-09-372-036-31

Perfect score: 10

Sequence: 1 PVAPQTQEVKK 10

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 3425486 residues

Word size: 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database: SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	100.0	484	PE0_LISMO	P21171 listeria mo
2	60.0	617	ABP1_SACEX	P38479 saccharomyc
3	60.0	985	ENV_FOAMV	P14351 human spuma
4	60.0	1133	ATX3_TETTH	Q95050 tetrahymena
5	50.0	102	HCC2_CRVCO	O01238 cryptosporidi
6	50.0	103	KAC4_RABIT	P01840 oryctolagus
7	50.0	153	PR51_SOLTU	P17641 solanum tub
8	50.0	174	RPC1_BPMU	P06019 bacterioph
9	50.0	185	EEPI_CHLPN	Q92900 chlamydia p
10	50.0	187	VGG_BPPHK	Q38042 bacterioph
11	50.0	200	RL29_MYCGE	P47405 mycoplasma
12	50.0	201	IMMU_BPSPB	P06650 mycoplasma
13	50.0	206	VAD1_TREPA	O83443 bacterioph
14	50.0	209	VG40_BPPH8	P14817 treponema p
15	50.0	217	V33P_ADEA1	P19416 human adeno
16	50.0	217	YD67_SCHPO	Q10319 schizosacch
17	50.0	228	TRKA_PYRHO	O57719 pyrococcus
18	50.0	239	KLK2_CAVPO	P12323 cavia porce
19	50.0	246	Y984_CAMEE	P45491 campylobact
20	50.0	250	YGB7_MYCPN	Q50313 mycoplasma
21	50.0	251	CB24_ARATH	P21521 arabidopsis
22	50.0	270	URED_SYNY3	P73047 synechocyst
23	50.0	296	YB27_YEAST	P38279 saccharomyc
24	50.0	305	LXD2_PHOLE	Q06878 photobacter
25	50.0	309	POLG_HCVH7	P27955 hepatitis c
26	50.0	309	URIG_SOYBN	P47663 glycine max
27	50.0	309	URID_SOYBN	O61672 m. equilibra
28	50.0	327	ENT2_MOUSE	P21574 xenopus lae
29	50.0	336	YB56_XENLA	P42390 zea mays (m
30	50.0	346	YMA3_MAIZE	Q02279 mycobacteri
31	50.0	354	YMA3_MYCBO	P34889 caenorhabdi
32	50.0	360	WMT2_CAEBL	P03526 reovirus (t
33	50.0	366	RRPO_REOVD	

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/132,653  
FILING DATE: 05-OCT-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 42 33 646.5  
FILING DATE: 06-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 42 35 718.7  
FILING DATE: 22-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 42 44 541.8  
FILING DATE: 30-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 43 18 186.4  
FILING DATE: 01-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Carol P. Einaudi  
REGISTRATION NUMBER: 32,220  
REFERENCE/DOCKET NUMBER: 05495-0001-03000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 59:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 498 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
US-08-471-770-59

Query Match 50.0%; Score 5; DB 1; Length 498;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 TOEVK 9  
|||||  
Db 309 TOEVK 313

Search completed: August 15, 2001, 12:33:54  
Job time: 59 sec

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,202  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/132,653  
FILING DATE: 05-OCT-1993  
APPLICATION NUMBER: DE P 42 33 646.5  
FILING DATE: 06-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 42 35 718.7  
FILING DATE: 22-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 42 44 541.8  
FILING DATE: 30-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 43 18 186.4  
FILING DATE: 01-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael J. Blake  
REGISTRATION NUMBER: 37,096  
REFERENCE/DOCKET NUMBER: 05495-0001-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 59:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 498 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-470-202-59

Query Match 50.0%; Score 5; DB 1; Length 498;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TOEVK 9  
|||||  
DB 309 TOEVK 313

RESULT 49  
US-08-470-202-60  
Sequence 60, Application US/08470202  
Patent No. 5759808  
GENERAL INFORMATION:  
APPLICANT: Guertler, Lutz G.  
APPLICANT: Eberle, Josef  
APPLICANT: Brunn, Albrecht V.  
APPLICANT: Knapp, Stefan  
APPLICANT: Hauser, Hans-Peter  
TITLE OF INVENTION: Retrovirus from the HIV Group and Its  
TITLE OF INVENTION: Use  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,202  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/132,653  
FILING DATE: 05-OCT-1993  
APPLICATION NUMBER: DE P 42 33 646.5  
FILING DATE: 06-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 42 35 718.7  
FILING DATE: 22-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 42 44 541.8  
FILING DATE: 30-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE P 43 18 186.4  
FILING DATE: 01-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Michael J. Blake  
REGISTRATION NUMBER: 37,096  
REFERENCE/DOCKET NUMBER: 05495-0001-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-408-4000  
TELEFAX: 202-408-4400  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 498 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-470-202-60

Query Match 50.0%; Score 5; DB 1; Length 498;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TOEVK 9  
|||||  
DB 309 TOEVK 313

RESULT 50  
US-08-471-770-59  
Sequence 59, Application US/08471770  
Patent No. 5770427  
GENERAL INFORMATION:  
APPLICANT: Guertler, Lutz G.  
APPLICANT: Eberle, Josef  
APPLICANT: Brunn, Albrecht V.  
APPLICANT: Knapp, Stefan  
APPLICANT: Hauser, Hans-Peter  
TITLE OF INVENTION: Retrovirus from the HIV Group and Its  
TITLE OF INVENTION: Use  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flinnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,770  
FILING DATE: 06-JUN-1995

Query Match 50.0%; Score 5; DB 2; Length 469;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
Db 361 PVAPT 365

## RESULT 45

US-09-071-709-12  
Sequence 12, Application US/09071709  
Patent No. 6171790  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Tang, Y. Tom  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: HUMAN PROTEASE ASSOCIATED PROTEINS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,709  
FILING DATE: Filed Herewith  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: CERRONE, MICHAEL C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0513 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 470 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-071-709-12

Query Match 50.0%; Score 5; DB 4; Length 470;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEYK 9  
Db 33 TOEYK 37

## RESULT 46

US-08-889-841-8  
Sequence 8, Application US/08889841B  
Patent No. 6090392  
GENERAL INFORMATION:  
APPLICANT: Berman, Phillip W.  
TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE  
FILE REFERENCE: 14918-703CIP  
CURRENT APPLICATION NUMBER: US/08/889,841B  
CURRENT FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: US 60/676,737  
EARLIER FILING DATE: 1996-07-08  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 486  
TYPE: PRT  
ORGANISM: HIV  
US-08-889-841-8

OY 2 VAPTO 6  
Db 457 VAPTO 461

Query Match 50.0%; Score 5; DB 3; Length 486;  
Best Local Similarity 100.0%; Pred. No. 3.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAPTO 6  
Db 457 VAPTO 461

## RESULT 47

US-08-889-841-10  
Sequence 10, Application US/08889841B  
Patent No. 6090392  
GENERAL INFORMATION:  
APPLICANT: Berman, Phillip W.  
TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE  
FILE REFERENCE: 14918-703CIP  
CURRENT APPLICATION NUMBER: US/08/889,841B  
CURRENT FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: US 60/676,737  
EARLIER FILING DATE: 1996-07-08  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 491  
TYPE: PRT  
ORGANISM: HIV  
US-08-889-841-10

Query Match 50.0%; Score 5; DB 3; Length 491;  
Best Local Similarity 100.0%; Pred. No. 3.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAPTO 6  
Db 457 VAPTO 461

## RESULT 48

US-08-470-202-59  
Sequence 59, Application US/08470202  
Patent No. 5759808  
GENERAL INFORMATION:  
APPLICANT: Guertler, Lutz G.  
APPLICANT: Eberle, Josef  
APPLICANT: Brunn, Albrecht V.  
APPLICANT: Knapp, Stefan  
APPLICANT: Hauser, Hans-Peter  
TITLE OF INVENTION: Retrovirus from the HIV Group and Its  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESS: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

APPLICANT: Okada, Shigenori  
APPLICANT: Adachi, Kenichi  
TITLE OF INVENTION: BILE ACID SULFATASE GENE, PLASMID  
TITLE OF INVENTION: CONTAINING SAID GENE AND METHOD OF PRODUCING BILE ACID  
TITLE OF INVENTION: SULFATE SULFATASE  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobb, Martens, Olson and Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/140,104A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP PCT/JP93/00244  
FILING DATE: 26-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Altman, Daniel E.  
REGISTRATION NUMBER: 34,115  
REFERENCE/DOCKET NUMBER: SAGS03.001AUS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 714-760-0404  
TELEFAX: 714-760-9502  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-140-104A-2

Query Match 50.0%; Score 5; DB 1; Length 467;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PTOEV 8  
DB 431 PTOEV 435

RESULT 43  
US-08-353-550-6  
Sequence 6, Application US/08353550  
Patent No. 5744313  
GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Cavanaugh, William M.  
TITLE OF INVENTION: Tyrosine Phosphorylated Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourile and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/353,550  
FILING DATE: 09-DEC-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 02307K-057300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ. ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 469 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-353-550-6

Query Match 50.0%; Score 5; DB 1; Length 469;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5  
DB 361 PVAPT 365

RESULT 44  
US-08-551-687-6  
Sequence 6, Application US/08551687  
Patent No. 5925547  
GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Cavanaugh, William M.  
TITLE OF INVENTION: Tyrosine Phosphorylated Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/551,687  
FILING DATE: 01-NOV-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/353,550  
FILING DATE: 09-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Matthew B.  
REGISTRATION NUMBER: P39,787  
REFERENCE/DOCKET NUMBER: 2307K-5731  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ. ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 469 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-551-687-6



CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/353,550  
FILING DATE: 09-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Matthew B.  
REGISTRATION NUMBER: P39,787  
REFERENCE/DOCKET NUMBER: 2307K-5731  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 428 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-551-687-1

Query Match 50.0%; Score 5; DB 2; Length 428;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
11111  
Db 361 PVAPT 365

RESULT 40  
US-08-807-342B-5  
Sequence 5, Application US/08807342B  
Patent No. 6077686  
GENERAL INFORMATION:  
APPLICANT: Def, Channing  
APPLICANT: O'Bryan, John P.  
APPLICANT: Pawson, Anthony  
TITLE OF INVENTION: No. 6077686el SHC Proteins  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/807,342B  
FILING DATE: 28-FEB-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/012,516  
FILING DATE: 29-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: MTS2USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 431 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein

US-08-807-342B-5

Query Match 50.0%; Score 5; DB 3; Length 431;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
11111  
Db 364 PVAPT 368

RESULT 41  
US-08-679-635A-3  
Sequence 3, Application US/08679635A  
Patent No. 5985643  
GENERAL INFORMATION:  
APPLICANT: Tomasz, Alexander  
APPLICANT: Delencastre, Herminia  
TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF  
METHICILLIN RESISTANT BACTERIA AND ANTAGONISTS THEREOF  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
FLOOR  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/679,635A  
FILING DATE: 10-JUL-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-141  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 445 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-679-635A-3

Query Match 50.0%; Score 5; DB 2; Length 445;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 OEYK 10  
11111  
Db 228 OEYK 232

RESULT 42  
US-08-140-104A-2  
Sequence 2, Application US/08140104A  
Patent No. 5585255  
GENERAL INFORMATION:  
APPLICANT: Tsukada, Yogi  
APPLICANT: Tazuke, Yasuhiko

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-839-581A-2

Query Match 50.0%; Score 5; DB 2; Length 399;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTOE 7  
IIIII  
DB 357 APTOE 361

RESULT 37  
US-09-023-591A-2  
Sequence 2, Application US/09023591A  
Patent No. 6210914  
GENERAL INFORMATION:  
APPLICANT: Staunton, Donald E.  
APPLICANT: Lipsey, Brian P.  
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin  
TITLE OF INVENTION: Binding/Signalling  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker, 6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/023,591A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,581  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Young J. Suh  
REGISTRATION NUMBER: P-41,337  
REFERENCE/DOCKET NUMBER: 27866/33886  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
INFORMATION FOR SEQ ID NO. 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 399 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-023-591A-2

Query Match 50.0%; Score 5; DB 4; Length 399;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTOE 7  
IIIII  
DB 357 APTOE 361

RESULT 38  
US-08-353-550-1  
Sequence 1, Application US/08353550  
Patent No. 5744313

GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Cavanaugh, William M.  
TITLE OF INVENTION: No. 5744313el Protein Domain Which Binds  
TITLE OF INVENTION: Tyrosine Phosphorylated Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESS: Townsend and Townsend Kourie and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/353,550  
FILING DATE: 09-DEC-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 02307K-057300  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO. 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 428 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-353-550-1

Query Match 50.0%; Score 5; DB 1; Length 428;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5  
IIIII  
DB 361 PVAPT 365

RESULT 39  
US-08-551-687-1  
Sequence 1, Application US/08551687  
Patent No. 5925547  
GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Cavanaugh, William M.  
TITLE OF INVENTION: No. 5925547el Protein Domain Which Binds  
TITLE OF INVENTION: Tyrosine Phosphorylated Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESS: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/551,687  
FILING DATE: 01-NOV-1995

STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,504B  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/459,595  
FILING DATE: 02-JUN-1995  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8587  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-459-504B-19

Query Match 50.0%; Score 5; DB 3; Length 346;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 OEYK 10  
DB 278 OEYK 282

RESULT 35  
US-08-459-444-19  
Sequence 19, Application US/08459444A  
Patent No. 6121014  
GENERAL INFORMATION:  
APPLICANT: Kozziel, Michael G.  
Desai, Nalin M.  
Lewis, Kelly S.  
Kramer, Vance C.  
Warren, Gregory W.  
Evoila, Stephen V.  
Crossland, Lyle D.  
Wright, Martha S.  
Merlin, Ellis J.  
Launis, Karen L.  
TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED  
NUCLEIC ACID CODING SEQUENCE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,444A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: S-18805/PL/CGC1577/CIP/DIV6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8587  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-08-459-444-19

Query Match 50.0%; Score 5; DB 3; Length 346;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 OEYK 10  
DB 278 OEYK 282

RESULT 36  
US-08-839-581A-2  
Sequence 2, Application US/08839581A  
Patent No. 5938705  
GENERAL INFORMATION:  
APPLICANT: Staunton, Donald E.  
Lipsky, Brian P.  
TITLE OF INVENTION: Cytoplasmic Modulators of Integrin  
Binding/Signaling  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker, 6300 Sears Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,581A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Young J. Suh  
REGISTRATION NUMBER: P-41,337  
REFERENCE/DOCKET NUMBER: 27866/33886  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 399 amino acids

STREET: Patent & Trademark Dept., 520 White Plains  
CITY: Tarrytown  
STATE: New York  
COUNTRY: USA  
ZIP: 10591-9005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,448A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40403  
REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8582  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-459-448A-19

Query Match 50.0%; Score 5; DB 2; Length 346;  
Best Local Similarity 100.0%; Pred. No. 2,6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 GEVKK 10  
|||||  
DB 278 GEVKK 282

RESULT 33  
US-08-459-595A-19  
Sequence 19, Application US/08459595A  
Patent No. 6018104  
GENERAL INFORMATION:  
APPLICANT: Kozziel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Ewola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttle, Janet L.  
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6018104artis Corporation  
STREET: Patent & Trademark Dept., 520 White Plains

STREET: Rd., POB 2005  
CITY: Tarrytown  
STATE: New York  
COUNTRY: USA  
ZIP: 10591-9005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,595A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40403  
REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8582  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-459-595A-19

Query Match 50.0%; Score 5; DB 3; Length 346;  
Best Local Similarity 100.0%; Pred. No. 2,6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 GEVKK 10  
|||||  
DB 278 GEVKK 282

RESULT 34  
US-08-459-504B-19  
Sequence 19, Application US/08459504B  
Patent No. 6075185  
GENERAL INFORMATION:  
APPLICANT: Kozziel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Ewola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttle, Janet L.  
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6075185artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park

```

: GENERAL INFORMATION:
: APPLICANT: Chien, David Y.
: APPLICANT: Ruten, William J.
: TITLE OF INVENTION: NABV Diagnostics and Vaccines
: NUMBER OF SEQUENCES: 777
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chiron Corporation
: STREET: 4560 Horton Street
: CITY: Emeryville
: STATE: CA
: COUNTRY: USA
: ZIP: 94608-2916
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/444,818
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/403,590
: FILING DATE: 14-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Harbin, Alisa A.
: REGISTRATION NUMBER: 33,895
: REFERENCE/DOCKET NUMBER: 0110.002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (508)359-3876
: TELEFAX: (508)359-3885
: INFORMATION FOR SEQ ID NO: 156:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 339 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: ORIGINAL SOURCE:
: INDIVIDUAL ISOLATE: Human 27
:
: US-08-444-818-156
:
: Query Match 50.0%; Score 5; DB 4; Length 339;
: Best Local Similarity 100.0%; Pred. No. 2.5e+02;
: Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 1 PVAFT 5
: 11111
: Db 113 PVAFT 117
:
: RESULT 31
: US-07-951-715A-19
: Sequence 19, Application US/07951715A
: Patent No. 5625136
: GENERAL INFORMATION:
: APPLICANT: Koziel, Michael G.
: APPLICANT: Desai, Nalini M.
: APPLICANT: Lewis, Kelly S.
: APPLICANT: Kramer, Vance C.
: APPLICANT: Warren, Gregory W.
: APPLICANT: Evola, Stephen V.
: APPLICANT: Crossland, Lyle D.
: APPLICANT: Wright, Martha S.
: APPLICANT: Merlin, Ellis J.
: APPLICANT: Launis, Karen L.
: APPLICANT: Rothstein, Steven J.
: APPLICANT: Bowman, Cindy G.
: APPLICANT: Dawson, John L.
: APPLICANT: Dunder, Erik M.
: APPLICANT: Pace, Gary M.
: APPLICANT: Suttie, Janet L.

```

```

: TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
: TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
: NUMBER OF SEQUENCES: 94
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: CIBA-GEIGY Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: New York
: COUNTRY: USA
: ZIP: 10532
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30B
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/951,715A
: FILING DATE: 25-SEP-1992
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/772,027
: FILING DATE: 04-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Spruill, W. Murray
: REGISTRATION NUMBER: 32,943
: REFERENCE/DOCKET NUMBER: S-18805/A/CCG 1577/CIP
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919)541-8615
: TELEFAX: (919)541-8689
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 346 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-07-951-715A-19
:
: Query Match 50.0%; Score 5; DB 1; Length 346;
: Best Local Similarity 100.0%; Pred. No. 2.6e+02;
: Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 6 QEVK 10
: 11111
: Db 278 QEVK 282
:
: RESULT 32
: US-08-459-448A-19
: Sequence 19, Application US/08459448A
: Patent No. 5859336
: GENERAL INFORMATION:
: APPLICANT: Koziel, Michael G.
: APPLICANT: Desai, Nalini M.
: APPLICANT: Lewis, Kelly S.
: APPLICANT: Kramer, Vance C.
: APPLICANT: Warren, Gregory W.
: APPLICANT: Evola, Stephen V.
: APPLICANT: Crossland, Lyle D.
: APPLICANT: Wright, Martha S.
: APPLICANT: Merlin, Ellis J.
: APPLICANT: Launis, Karen L.
: APPLICANT: Rothstein, Steven J.
: APPLICANT: Bowman, Cindy G.
: APPLICANT: Dawson, John L.
: APPLICANT: Dunder, Erik M.
: APPLICANT: Pace, Gary M.
: APPLICANT: Suttie, Janet L.
: TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED
: TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE
: NUMBER OF SEQUENCES: 94
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 5859336artis Corporation

```

SUSAN D., POPE, SHARON H., STRAUSBERG, SUSAN L., RUFF, MICHAEL D.,  
; AUGUSTINE, PATRICIA C., DANFORTH, HARRY D.  
; TITLE OF INVENTION: EIMERIA ANTIGENIC COMPOSITION WHICH  
; ELICITS ANTIBODIES AGAINST AVIAN COCCIDIOSIS  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/148,432  
; FILING DATE: 08-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 581,693  
; FILING DATE: 12-SEP-1990  
; APPLICATION NUMBER: 215,162  
; FILING DATE: 05-JUL-1989  
; APPLICATION NUMBER: 746,520  
; FILING DATE: 19-JUN-1985  
; APPLICATION NUMBER: 627,811  
; FILING DATE: 05-JUL-1984  
; SEQ ID NO: 2:  
; LENGTH: 242  
5482709-2

Query Match 50.0%; Score 5; DB 6; Length 242;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PTOEV 8  
; | | | | |  
Db 36 PTOEV 40

RESULT 28  
US-08-810-572A-2  
; Sequence 2, Application US/08810572A  
; Patent No. 5869102  
; GENERAL INFORMATION:  
; APPLICANT: Bram, Richard J.  
; APPLICANT: von Bulow, Gotz  
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS  
; TITLE OF INVENTION: CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE  
; NUMBER OF SEQUENCES: 11  
; FILING DATE: THEREOF  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/810,572A  
; FILING DATE: 28-FEB-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1340-1-007  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 293 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

HYPOTHETICAL: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
US-08-810-572A-2

Query Match 50.0%; Score 5; DB 2; Length 293;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APTOE 7  
; | | | | |  
Db 239 APTOE 243

RESULT 29  
US-08-444-818-152  
; Sequence 152, Application US/08444818  
; Patent No. 6150087  
; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.  
; APPLICANT: Ruter, William J.  
; TITLE OF INVENTION: NANV Diagnostics and Vaccines  
; NUMBER OF SEQUENCES: 777  
; CURRENT APPLICATION DATA:  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Ailsa A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110,002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876  
; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 152:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 339 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-444-818-152

Query Match 50.0%; Score 5; DB 4; Length 339;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
; | | | | |  
Db 113 PVAPT 117

RESULT 30  
US-08-444-818-156  
; Sequence 156, Application US/08444818  
; Patent No. 6150087

TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND  
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS  
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE  
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES  
NUMBER OF SEQUENCES: 263  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/10398  
FILING DATE: 15-AUG-1995  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/086,428  
FILING DATE: 29 JUNE 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/290/665  
FILING DATE: 15 AUGUST 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
ORGANISM: homosapiens  
INDIVIDUAL ISOLATE: S18  
PCT-US95-10398-57

Query Match 50.0%; Score 5; DB 5; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAET 5  
11111  
DB 50 PVAET 54

RESULT 25  
US-08-969-721-8  
Sequence 8, Application US/08969721  
Patent No. 5972339  
GENERAL INFORMATION:  
APPLICANT: Walker, Bruce D.  
TITLE OF INVENTION: METHODS OF ELICITING ANTI-HIV-1  
TITLE OF INVENTION: HELPER T CELL RESPONSES  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/969,721  
FILING DATE: 13-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Ph.D., J.D., Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 08472/732001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 231 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-969-721-8

Query Match 50.0%; Score 5; DB 2; Length 231;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TOEYK 9  
11111  
DB 178 TOEYK 182

RESULT 26  
5273901-3  
Patent No. 5273901  
APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,  
SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;  
AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.  
TITLE OF INVENTION: GENETICALLY ENGINEERED COCCIDIOSIS  
SPOROZOITE 21.5 KB ANTIGEN, AC-6B  
NUMBER OF SEQUENCES: 11  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/581,693  
FILING DATE: 12-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 215,162  
FILING DATE: 05-JUL-1988  
APPLICATION NUMBER: 746,520  
FILING DATE: 19-JUN-1985  
APPLICATION NUMBER: 627,811  
FILING DATE: 05-JUL-1984  
SEQ ID NO: 3  
LENGTH: 242  
5273901-3

Query Match 50.0%; Score 5; DB 6; Length 242;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PTOEV 8  
11111  
DB 36 PTOEV 40

RESULT 27  
5482709-2  
Patent No. 5482709  
APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
ORIGINAL SOURCE:  
ORGANISM: homosapiens  
INDIVIDUAL ISOLATE: S18  
US-08-468-570-57

Query Match 50.0%; Score 5; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5  
|1111  
DB 50 PVAPT 54

RESULT 22  
US-08-290-665A-57  
Sequence 57, Application US/08290665A  
Patent No. 5882852  
GENERAL INFORMATION:  
APPLICANT: BURKH, J., MILLER, R.H. AND  
APPLICANT: PURCELL, R.H.  
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND  
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS  
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE  
NUMBER OF SEQUENCES: 263  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290.665A  
FILING DATE: 15-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
ORIGINAL SOURCE:  
ORGANISM: homosapiens  
INDIVIDUAL ISOLATE: S18  
US-08-290-665A-57

Query Match 50.0%; Score 5; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5  
|1111  
DB 50 PVAPT 54

RESULT 23  
US-08-444-818-167  
Sequence 167, Application US/08444818  
Patent No. 6150087  
GENERAL INFORMATION:  
APPLICANT: Chien, David Y.  
APPLICANT: Rutler, William J.  
TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
NUMBER OF SEQUENCES: 777  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444.818  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/403.590  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Harbin, Ailsa A.  
REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885  
INFORMATION FOR SEQ ID NO: 167:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: JH27  
US-08-444-818-167

Query Match 50.0%; Score 5; DB 4; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5  
|1111  
DB 125 PVAPT 129

RESULT 24  
PCT-US95-10398-57  
Sequence 57, Application PC/TUS9510398  
GENERAL INFORMATION:  
APPLICANT: BURKH, J., MILLER, R.H. AND  
APPLICANT: PURCELL, R.H.  
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED



US-08-231-368-44  
; Sequence 44, Application US/08231368  
; Patent No. 5756312  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, Amy J.  
; APPLICANT: Houghton, Michael  
; TITLE OF INVENTION: Immunoreactive polypeptide Compositions  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,368  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/759,575  
; FILING DATE: 13-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0205.001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2708  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 192 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-231-368-44

Query Match 50.0%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5  
|||||  
Db 50 PVAPT 54

RESULT 20  
US-08-440-210-44  
; Sequence 44, Application US/08440210  
; Patent No. 5766845  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, Amy J.  
; APPLICANT: Houghton, Michael  
; TITLE OF INVENTION: Immunoreactive polypeptide Compositions  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,210  
; FILING DATE: 12-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,368  
; FILING DATE:  
; APPLICATION NUMBER: US 07/759,575  
; FILING DATE: 13-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0205.001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2708  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 192 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-440-210-44

Query Match 50.0%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5  
|||||  
Db 50 PVAPT 54

RESULT 21  
US-08-468-570-57  
; Sequence 57, Application US/08468570  
; Patent No. 5871962  
; GENERAL INFORMATION:  
; APPLICANT: BURK, J., MILLER, R.H. AND  
; APPLICANT: PURCELL, R.H.  
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE  
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE  
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN  
; NUMBER OF SEQUENCES: 159  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/468,570  
; FILING DATE: 6-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/086,428  
; FILING DATE: 29-JUN-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: RICHARD W. BORK  
; REGISTRATION NUMBER: 36,459  
; REFERENCE/DOCKET NUMBER: 2026-4070US1

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/086.428B  
;; FILING DATE: 29-JUN-1993  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: RICHARD W. BORK  
;; REGISTRATION NUMBER: 36,459  
;; REFERENCE/DOCKET NUMBER: 2026-4070  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 758-4800  
;; TELEFAX: (212) 751-6849  
;;  
;; TELETYPE: 421792  
;; INFORMATION FOR SEQ ID NO: 57:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 192 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: unknown  
;; TOPOLOGY: unknown  
;; ORIGINAL SOURCE:  
;; ORGANISM: homosapiens  
;; INDIVIDUAL ISOLATE: S18  
;;  
US-08-086-428B-57  
;  
Query Match 50.0%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
;  
QY 1 PVAPT 5  
Db 50 PVAPT 54  
;  
RESULT 17  
US-08-440-103-44  
;; Sequence 44, Application US/08440103  
;; Patent No. 5670152  
;; GENERAL INFORMATION:  
;; APPLICANT: Weiner, Amy J.  
;; APPLICANT: Houghton, Michael  
;; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions  
;; NUMBER OF SEQUENCES: 45  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Chiron Corporation  
;; STREET: 4560 Horton Street  
;; CITY: Emeryville  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94608  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/440.103  
;; FILING DATE: 12-MAY-1995  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/231.368  
;; FILING DATE:  
;; APPLICATION NUMBER: US/07/759.575  
;; FILING DATE: 13-SEP-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: McClung, Barbara G.  
;; REGISTRATION NUMBER: 33,113  
;; REFERENCE/DOCKET NUMBER: 0205.001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (510) 601-2708  
;; TELEFAX: (510) 655-3542  
;; INFORMATION FOR SEQ ID NO: 44:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 192 amino acids

;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-440-103-44  
;  
Query Match 50.0%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
;  
QY 1 PVAPT 5  
Db 50 PVAPT 54  
;  
RESULT 18  
US-08-440-542-44  
;; Sequence 44, Application US/08440542  
;; Patent No. 5670153  
;; GENERAL INFORMATION:  
;; APPLICANT: Weiner, Amy J.  
;; APPLICANT: Houghton, Michael  
;; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions  
;; NUMBER OF SEQUENCES: 45  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Chiron Corporation  
;; STREET: 4560 Horton Street  
;; CITY: Emeryville  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94608  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/440.542  
;; FILING DATE: 12-MAY-1995  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/231.368  
;; FILING DATE:  
;; APPLICATION NUMBER: US/07/759.575  
;; FILING DATE: 13-SEP-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: McClung, Barbara G.  
;; REGISTRATION NUMBER: 33,113  
;; REFERENCE/DOCKET NUMBER: 0205.001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (510) 601-2708  
;; TELEFAX: (510) 655-3542  
;; INFORMATION FOR SEQ ID NO: 44:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 192 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-440-542-44  
;  
Query Match 50.0%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
;  
QY 1 PVAPT 5  
Db 50 PVAPT 54  
;  
RESULT 19

RESULT 14  
US-08-810-572A-4  
; Sequence 4, Application US/08810572A  
; Patent No. 5969102  
; GENERAL INFORMATION:  
; APPLICANT: Bram, Richard J.  
; APPLICANT: von Bulow, Gotz  
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS  
; TITLE OF INVENTION: CAM, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE  
; TITLE OF INVENTION: THEREOF  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; STREET: Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/810,572A  
; FILING DATE: 28-FEB-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1340-1-007  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 107 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: C-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; US-08-810-572A-4

Query Match 50.0%; Score 5; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APTQE 7  
|1111  
Db .53 APTQE 57

RESULT 15  
US-08-497-025-10  
; Sequence 10, Application US/08497025  
; Patent No. 5646251  
; GENERAL INFORMATION:  
; APPLICANT: Ruegg, Curtis L.  
; APPLICANT: Rivas, Alberto  
; APPLICANT: Laus, Reiner  
; APPLICANT: Engleman, Edgar G.  
; TITLE OF INVENTION: ALLOREACTION-ASSOCIATED ANTIGEN (ARA9):  
; TITLE OF INVENTION: A NOVEL MEMBER OF THE IMMUNOGLOBULIN GENE SUPERFAMILY  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew

STREET: 379 Lytton Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: US  
ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/497,025  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/149,212  
; FILING DATE: 05-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M.  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 5490A-219  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 118 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..118  
; OTHER INFORMATION: /note="A variable-type  
; OTHER INFORMATION: Immunoglobulin-type domain of ARA9-h-1."  
; US-08-497-025-10

Query Match 50.0%; Score 5; DB 1; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PTOEV 8  
|1111  
Db 43 PTOEV 47

RESULT 16  
US-08-086-428B-57  
; Sequence 57, Application US/08086428B  
; Patent No. 5514539  
; GENERAL INFORMATION:  
; APPLICANT: BURK, J., MILLER, R.H. AND  
; APPLICANT: PURCELL, R.H.  
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE  
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE  
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN  
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES  
; NUMBER OF SEQUENCES: 159  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1

APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BOBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1-7  
OTHER INFORMATION: /note= "SOME OR ALL Xaa AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 14-20  
OTHER INFORMATION: /note= "SOME OR ALL Xaa AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
US-08-456-670B-17

Query Match 60.0%; Score 6; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 PVAPQ 6  
DB 8 PVAPQ 13

RESULT 13  
US-08-456-670B-35  
Sequence 35 Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKELEIR, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BOBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria innocua  
US-08-456-670B-35

Query Match 50.0%; Score 5; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 QEVK 10  
DB 7 QEVK 11

OY 4 PTOEVKK 10  
|||||  
Db 1 PTOEVKK 7

## RESULT 10

US-08-482-847-28  
; Sequence 28, Application US/08482847  
; Patent No. 5556757  
; GENERAL INFORMATION:  
; APPLICANT: VAN ALSTYNE, Diane  
; APPLICANT: SHARMA, Lawrence Rajendra  
; TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
; TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
; TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,847  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/127,499  
; FILING DATE: 28-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 51916/104/INBI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
US-08-482-847-28

Query Match 70.0%; Score 7; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PTOEVKK 10  
|||||  
Db 1 PTOEVKK 7

## RESULT 11

US-08-456-670B-26  
; Sequence 26, Application US/08456670B  
; Patent No. 5932415  
; GENERAL INFORMATION:  
; APPLICANT: SCHUBERT, PETER  
; APPLICANT: NEUMANN, Siegfried  
; APPLICANT: PAMELZIK, MARTINA  
; APPLICANT: LINXWETTER, WINFRIED  
; APPLICANT: BURGER, CHRISTA  
; APPLICANT: HOFMANN, GOTTFRIED

APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US

## ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,670B

FILING DATE: 01-JUN-1995

CLASSIFICATION: 435

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/412,227

FILING DATE: 27-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/075,248

FILING DATE: 11-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4239567.4

FILING DATE: 25-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4219111.4

FILING DATE: 11-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: HAMLET-KING, DIANA

REGISTRATION NUMBER: 33,302

REFERENCE/DOCKET NUMBER: MERCK 1694D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-243-6333

TELEFAX: 703-243-6410

TELEX: 64191

INFORMATION FOR SEQ ID NO: 26:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORGANISM: Listeria monocytogenes

STRAIN: EGD

US-08-456-670B-26

Query Match 60.0%; Score 6; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPQ 6  
|||||  
Db 4 PVAPQ 9

## RESULT 12

US-08-456-670B-17  
; Sequence 17, Application US/08456670B  
; Patent No. 5932415  
; GENERAL INFORMATION:  
; APPLICANT: SCHUBERT, PETER  
; APPLICANT: NEUMANN, Siegfried  
; APPLICANT: PAMELZIK, MARTINA  
; APPLICANT: LINXWETTER, WINFRIED  
; APPLICANT: BURGER, CHRISTA

APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-26

Query Match 100.0%; Score 10; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PVAPTOEVKK 10  
|||  
Db 148 PVAPTOEVKK 157

RESULT 8  
US-08-482-847-26  
Sequence 26, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-26

Query Match 100.0%; Score 10; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PVAPTOEVKK 10  
|||  
Db 148 PVAPTOEVKK 157

RESULT 9  
US-08-127-499A-28  
Sequence 28, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-28

Query Match 70.0%; Score 7; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

APPLICANT: LINXWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ. ID NO.: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 232 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-39

Query Match 100.0%; Score 10; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 0.0018;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPTOEVRK 10  
| | | | | | | | | |  
Db 79 PVAPTOEVRK 88

RESULT 6  
US-08-456-670B-40  
Sequence 40, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED

APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINXWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ. ID NO.: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 478 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-40

Query Match 100.0%; Score 10; DB 2; Length 478;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPTOEVRK 10  
| | | | | | | | | |  
Db 148 PVAPTOEVRK 157

RESULT 7  
US-08-127-499A-26  
Sequence 26, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane

APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: *Listeria monocytogenes*  
STRAIN: EGD  
US-08-456-670B-30

Query Match 100.0%; Score 10; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PVAPTOEVKK 10  
Db 3 PVAPTOEVKK 12

RESULT 3  
US-08-127-499A-27  
Sequence 27, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-27

Query Match 100.0%; Score 10; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.00023;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PVAPTOEVKK 10  
Db 5 PVAPTOEVKK 14

RESULT 4  
US-08-482-847-27  
Sequence 27, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THERETO, AND USES THEREOF  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-27

Query Match 100.0%; Score 10; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.00023;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PVAPTOEVKK 10  
Db 5 PVAPTOEVKK 14

RESULT 5  
US-08-456-670B-39  
Sequence 39, Application US/08456670B  
Patent No. 5532415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA



977 4 40.0 300 3 US-08-705-875A-6 Sequence 6, Appl1  
978 4 40.0 300 4 US-09-220-731-21 Sequence 21, Appl1  
979 4 40.0 300 4 US-08-439-992A-5 Sequence 5, Appl1  
980 4 40.0 301 3 US-09-002-298-3 Sequence 3, Appl1  
981 4 40.0 302 1 US-07-640-029-6 Sequence 6, Appl1  
982 4 40.0 302 1 US-07-921-807B-7 Sequence 7, Appl1  
983 4 40.0 302 1 US-07-921-807B-8 Sequence 8, Appl1  
984 4 40.0 302 1 US-08-441-944A-7 Sequence 7, Appl1  
985 4 40.0 302 4 US-08-439-992A-6 Sequence 8, Appl1  
986 4 40.0 303 1 US-07-917-111-2 Sequence 2, Appl1  
987 4 40.0 303 1 US-07-917-111-3 Sequence 2, Appl1  
988 4 40.0 303 1 US-08-479-638-2 Sequence 3, Appl1  
989 4 40.0 303 1 US-08-479-638-3 Sequence 3, Appl1  
990 4 40.0 303 2 US-08-294-871A-2 Sequence 2, Appl1  
991 4 40.0 303 2 US-08-294-871A-4 Sequence 4, Appl1  
992 4 40.0 303 2 US-08-294-871A-6 Sequence 6, Appl1  
993 4 40.0 303 2 US-08-294-871A-8 Sequence 8, Appl1  
994 4 40.0 303 2 US-08-294-871A-10 Sequence 10, Appl1  
995 4 40.0 303 2 US-08-294-871A-12 Sequence 12, Appl1  
996 4 40.0 303 2 US-08-294-871A-14 Sequence 14, Appl1  
997 4 40.0 303 2 US-08-294-871A-16 Sequence 16, Appl1  
998 4 40.0 303 2 US-08-294-871A-18 Sequence 18, Appl1  
999 4 40.0 303 2 US-08-294-871A-18 Sequence 18, Appl1  
1000 4 40.0 303 2 US-08-294-871A-18 Sequence 20, Appl1

## ALIGNMENTS

RESULT 1  
US-08-456-670B-31  
Sequence 31, Application US/08456670B

Patent No. 5932415

GENERAL INFORMATION:

APPLICANT: SCHUBERT, PETER

APPLICANT: NEUMANN, SIEGFRIED

APPLICANT: PAMELZIK, MARTINA

APPLICANT: LINXWEILER, WINFRIED

APPLICANT: BURGER, CHRISTA

APPLICANT: HOFMANN, GOTTFRIED

APPLICANT: BOBERT, ANDREAS

APPLICANT: GOEBEL, WERNER

APPLICANT: KOHLER, STEFAN

TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.

STREET: 2200 CLARENDON BLVD., SUITE 1400

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: US

ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,670B

FILING DATE: 01-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/412,227

FILING DATE: 27-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/075,248

FILING DATE: 11-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4239567.4

FILING DATE: 25-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 421911.4

FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-31

Query Match 100.0%; Score 10; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00012;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPTOEVKK 10  
DB 1 PVAPTOEVKK 10

RESULT 2

US-08-456-670B-30

Sequence 30, Application US/08456670B

Patent No. 5932415

GENERAL INFORMATION:

APPLICANT: SCHUBERT, PETER

APPLICANT: NEUMANN, SIEGFRIED

APPLICANT: PAMELZIK, MARTINA

APPLICANT: LINXWEILER, WINFRIED

APPLICANT: BURGER, CHRISTA

APPLICANT: HOFMANN, GOTTFRIED

APPLICANT: BOBERT, ANDREAS

APPLICANT: GOEBEL, WERNER

APPLICANT: KOHLER, STEFAN

TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.

STREET: 2200 CLARENDON BLVD., SUITE 1400

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: US

ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,670B

FILING DATE: 01-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/412,227

FILING DATE: 27-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/075,248

FILING DATE: 11-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4239567.4

FILING DATE: 25-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 421911.4

831	4	40.0	232	2	US-08-615-271-8	Sequence 8, Appl	904	4	40.0	255	3	US-08-841-089-106	Sequence 106, App
832	4	40.0	232	3	US-09-074-660-8	Sequence 8, Appl	905	4	40.0	255	5	PCT-US95-04570-106	Sequence 106, App
833	4	40.0	232	3	US-09-074-659-8	Sequence 8, Appl	906	4	40.0	255	5	PCT-US95-04589-106	Sequence 106, App
834	4	40.0	232	3	US-09-106-468-8	Sequence 8, Appl	907	4	40.0	255	5	PCT-US96-01314-57	Sequence 57, Appl
835	4	40.0	232	4	US-09-106-466A-8	Sequence 8, Appl	908	4	40.0	257	6	5204259-5	Patent No. 5204259
836	4	40.0	232	4	US-09-106-467-8	Sequence 8, Appl	909	4	40.0	258	1	US-08-744-026-3	Sequence 3, Appl
837	4	40.0	232	6	5204259-1	Patent No. 5204259	910	4	40.0	258	2	US-09-102-732-3	Sequence 3, Appl
838	4	40.0	233	6	US-08-523-894-6	Sequence 6, Appl	911	4	40.0	258	4	US-08-961-083-00	Sequence 90, Appl
839	4	40.0	234	1	US-08-287-959-10	Sequence 10, Appl	912	4	40.0	258	4	US-09-261-767-3	Sequence 3, Appl
840	4	40.0	234	1	US-08-487-550-2	Sequence 2, Appl	913	4	40.0	258	6	5204259-3	Patent No. 5204259
841	4	40.0	235	1	US-08-458-516-23	Sequence 23, Appl	914	4	40.0	259	4	US-09-216-295-5	Sequence 5, Appl
842	4	40.0	235	2	US-08-378-939-12	Sequence 12, Appl	915	4	40.0	261	1	US-08-744-026-5	Sequence 5, Appl
843	4	40.0	235	4	US-09-049-672A-10	Sequence 10, Appl	916	4	40.0	261	2	US-09-102-732-5	Sequence 5, Appl
844	4	40.0	235	4	US-09-049-672A-12	Sequence 12, Appl	917	4	40.0	261	3	US-08-768-859A-6	Sequence 6, Appl
845	4	40.0	236	3	US-08-487-550-10	Sequence 10, Appl	918	4	40.0	261	3	US-08-768-859A-19	Sequence 19, Appl
846	4	40.0	236	3	US-08-493-071-15	Sequence 15, Appl	919	4	40.0	261	3	US-09-083-521-6	Sequence 6, Appl
847	4	40.0	236	4	US-09-049-672A-7	Sequence 7, Appl	920	4	40.0	261	3	US-08-767-820A-6	Sequence 6, Appl
848	4	40.0	236	4	US-08-961-083-96	Sequence 96, Appl	921	4	40.0	261	3	US-08-767-820A-19	Sequence 19, Appl
849	4	40.0	236	4	US-08-096-946-10	Sequence 10, Appl	922	4	40.0	261	3	US-08-622-046B-3	Sequence 3, Appl
850	4	40.0	237	1	US-08-096-946-10	Sequence 10, Appl	923	4	40.0	261	3	US-08-622-046B-14	Sequence 14, Appl
851	4	40.0	237	1	US-08-096-946-11	Sequence 11, Appl	924	4	40.0	261	4	US-09-261-767-5	Sequence 5, Appl
852	4	40.0	237	2	US-08-672-564-3	Sequence 3, Appl	925	4	40.0	261	4	US-09-100-264-7	Sequence 7, Appl
853	4	40.0	237	2	US-08-844-024-2	Sequence 2, Appl	926	4	40.0	261	5	PCT-US95-06157-6	Sequence 6, Appl
854	4	40.0	237	2	US-08-718-547-2	Sequence 2, Appl	927	4	40.0	262	1	US-08-106-981-6	Sequence 6, Appl
855	4	40.0	237	2	US-08-768-859A-1	Sequence 1, Appl	928	4	40.0	262	1	US-08-744-026-4	Sequence 4, Appl
856	4	40.0	237	3	US-08-768-859A-16	Sequence 16, Appl	929	4	40.0	262	1	US-08-790-137-1	Sequence 1, Appl
857	4	40.0	237	3	US-08-768-859A-21	Sequence 21, Appl	930	4	40.0	262	2	US-08-790-137-3	Sequence 3, Appl
858	4	40.0	237	3	US-08-767-820A-1	Sequence 1, Appl	931	4	40.0	262	2	US-08-681-151-4	Sequence 4, Appl
859	4	40.0	237	3	US-08-767-820A-16	Sequence 16, Appl	932	4	40.0	262	2	US-09-102-732-4	Sequence 4, Appl
860	4	40.0	237	3	US-08-767-820A-21	Sequence 21, Appl	933	4	40.0	262	2	US-08-824-874-4	Sequence 4, Appl
861	4	40.0	237	3	US-08-622-046B-1	Sequence 1, Appl	934	4	40.0	262	3	US-08-807-151-4	Sequence 4, Appl
862	4	40.0	237	3	US-08-622-046B-7	Sequence 7, Appl	935	4	40.0	262	3	US-09-025-059-4	Sequence 4, Appl
863	4	40.0	237	3	US-08-622-046B-12	Sequence 12, Appl	936	4	40.0	262	4	US-09-261-767-4	Sequence 4, Appl
864	4	40.0	237	4	US-08-944-483-37	Sequence 37, Appl	937	4	40.0	262	4	US-09-210-084-4	Sequence 4, Appl
865	4	40.0	237	4	US-08-944-483-38	Sequence 38, Appl	938	4	40.0	266	2	US-07-808-457-3	Sequence 3, Appl
866	4	40.0	237	4	US-09-100-264-1	Sequence 1, Appl	939	4	40.0	266	2	US-07-808-457-5	Sequence 5, Appl
867	4	40.0	237	4	US-09-100-264-3	Sequence 3, Appl	940	4	40.0	266	2	US-08-845-161A-33	Sequence 33, Appl
868	4	40.0	237	4	US-09-100-264-12	Sequence 12, Appl	941	4	40.0	266	4	US-09-270-751-33	Sequence 33, Appl
869	4	40.0	237	5	PCT-US94-07329-10	Sequence 10, Appl	942	4	40.0	266	5	PCT-US92-10178-3	Sequence 3, Appl
870	4	40.0	237	5	PCT-US94-07329-11	Sequence 11, Appl	943	4	40.0	266	5	PCT-US92-10178-5	Sequence 5, Appl
871	4	40.0	237	5	PCT-US95-06157-1	Sequence 1, Appl	944	4	40.0	269	4	US-08-646-265A-109	Sequence 109, App
872	4	40.0	237	5	PCT-US95-06157-16	Sequence 16, Appl	945	4	40.0	270	4	US-09-083-305-16	Sequence 16, Appl
873	4	40.0	238	3	US-08-768-859A-8	Sequence 8, Appl	946	4	40.0	272	2	US-08-160-524A-7	Sequence 7, Appl
874	4	40.0	238	3	US-08-767-820A-8	Sequence 8, Appl	947	4	40.0	272	2	US-08-492-027A-4	Sequence 4, Appl
875	4	40.0	238	4	US-08-944-483-39	Sequence 39, Appl	948	4	40.0	275	1	US-08-589-446-4	Sequence 4, Appl
876	4	40.0	238	4	US-09-216-295-12	Sequence 12, Appl	949	4	40.0	275	1	US-08-444-882-4	Sequence 4, Appl
877	4	40.0	238	5	PCT-US95-06157-8	Sequence 8, Appl	950	4	40.0	275	2	US-08-389-459A-4	Sequence 4, Appl
878	4	40.0	239	4	US-09-306-881-4	Sequence 4, Appl	951	4	40.0	275	3	US-08-987-867A-4	Sequence 4, Appl
879	4	40.0	240	1	US-08-472-228A-1	Sequence 1, Appl	952	4	40.0	277	1	US-08-690-457-5	Sequence 5, Appl
880	4	40.0	240	4	US-09-049-672A-11	Sequence 11, Appl	953	4	40.0	277	2	US-08-628-187-5	Sequence 5, Appl
881	4	40.0	240	5	PCT-US96-09303-1	Sequence 1, Appl	954	4	40.0	277	3	US-08-493-071-3	Sequence 3, Appl
882	4	40.0	240	6	5204445-2	Patent No. 5204445	955	4	40.0	278	1	US-07-921-807B-10	Sequence 10, Appl
883	4	40.0	244	3	US-08-768-859A-10	Sequence 10, Appl	956	4	40.0	278	1	US-08-188-382-9	Sequence 9, Appl
884	4	40.0	244	3	US-08-767-820A-10	Sequence 10, Appl	957	4	40.0	278	1	US-08-646-715-9	Sequence 9, Appl
885	4	40.0	244	3	US-08-622-046B-5	Sequence 5, Appl	958	4	40.0	278	1	US-08-341-944A-10	Sequence 10, Appl
886	4	40.0	244	3	US-08-622-046B-16	Sequence 16, Appl	959	4	40.0	279	3	US-08-397-411-13	Sequence 13, Appl
887	4	40.0	244	4	US-09-100-264-5	Sequence 5, Appl	960	4	40.0	282	5	PCT-US95-13335-1	Sequence 1, Appl
888	4	40.0	244	5	PCT-US95-06157-10	Sequence 10, Appl	961	4	40.0	283	6	5204259-9	Patent No. 5204259
889	4	40.0	245	2	US-08-912-129A-48	Sequence 48, Appl	962	4	40.0	284	2	US-08-767-096-3	Sequence 3, Appl
890	4	40.0	247	2	US-08-851-974-1	Sequence 1, Appl	963	4	40.0	286	1	US-08-382-184-3	Sequence 3, Appl
891	4	40.0	247	2	US-08-851-974-4	Sequence 4, Appl	964	4	40.0	286	2	US-08-641-356-3	Sequence 3, Appl
892	4	40.0	247	2	US-09-213-390-1	Sequence 1, Appl	965	4	40.0	286	4	US-09-132-528-4	Sequence 4, Appl
893	4	40.0	247	2	US-09-213-390-4	Sequence 4, Appl	966	4	40.0	286	4	US-08-875-484-3	Sequence 3, Appl
894	4	40.0	249	2	US-08-991-946A-1	Sequence 1, Appl	967	4	40.0	287	1	US-08-690-457-4	Sequence 4, Appl
895	4	40.0	251	4	US-09-446-504-1	Sequence 1, Appl	968	4	40.0	287	2	US-08-628-187-4	Sequence 2, Appl
896	4	40.0	251	4	US-08-944-483-28	Sequence 28, Appl	969	4	40.0	287	3	US-08-493-071-2	Sequence 2, Appl
897	4	40.0	252	1	US-08-253-877C-6	Sequence 6, Appl	970	4	40.0	288	1	US-08-690-457-3	Sequence 3, Appl
898	4	40.0	252	2	US-08-452-164A-6	Sequence 6, Appl	971	4	40.0	288	2	US-08-628-187-3	Sequence 3, Appl
899	4	40.0	252	2	US-08-824-405-4	Sequence 4, Appl	972	4	40.0	288	3	US-08-493-071-1	Sequence 1, Appl
900	4	40.0	253	1	US-08-622-353-8	Sequence 8, Appl	973	4	40.0	299	2	US-08-923-856-1	Sequence 1, Appl
901	4	40.0	253	1	US-08-622-352A-10	Sequence 10, Appl	974	4	40.0	299	3	US-09-216-294-1	Sequence 1, Appl
902	4	40.0	255	2	US-08-310-912A-106	Sequence 106, App	975	4	40.0	300	1	US-07-640-029-5	Sequence 5, Appl
903	4	40.0	255	3	US-08-826-390-10	Sequence 10, Appl	976	4	40.0	300	3	US-08-705-875A-4	Sequence 4, Appl

685	4	40.0	140	4	US-08-569-147-82	Sequence 82, Appl	758	4	40.0	185	2	US-08-865-336-3	Sequence 3, Appl1
686	4	40.0	140	4	US-08-579-378A-13	Sequence 12, Appl	759	4	40.0	186	1	US-08-928-443-3	Sequence 3, Appl1
687	4	40.0	140	5	PCT-US93-11612-12	Sequence 12, Appl	760	4	40.0	186	3	US-09-129-055-3	Sequence 3, Appl1
688	4	40.0	141	1	US-08-461-289-4	Sequence 4, Appl1	761	4	40.0	188	3	US-09-122-443-14	Sequence 14, Appl
689	4	40.0	141	1	US-08-462-939-4	Sequence 4, Appl1	762	4	40.0	189	4	US-08-026-758-18	Sequence 18, Appl
690	4	40.0	141	1	US-08-453-877C-4	Sequence 4, Appl1	763	4	40.0	190	4	US-09-271-970-14	Sequence 14, Appl
691	4	40.0	141	2	US-08-452-164A-4	Sequence 4, Appl1	764	4	40.0	192	1	US-08-086-428B-53	Sequence 53, Appl
692	4	40.0	142	2	US-08-561-521-17	Sequence 17, Appl	765	4	40.0	192	1	US-08-086-428B-58	Sequence 58, Appl
693	4	40.0	142	2	US-08-476-176B-12	Sequence 12, Appl	766	4	40.0	192	1	US-08-086-428B-85	Sequence 85, Appl
694	4	40.0	142	2	US-08-476-176B-14	Sequence 14, Appl	767	4	40.0	192	2	US-08-468-570-5	Sequence 5, Appl
695	4	40.0	142	2	US-08-476-176B-16	Sequence 16, Appl	768	4	40.0	192	2	US-08-468-570-58	Sequence 58, Appl
696	4	40.0	142	2	US-08-476-176B-18	Sequence 18, Appl	769	4	40.0	192	2	US-08-468-570-85	Sequence 85, Appl
697	4	40.0	142	3	US-08-127-721A-12	Sequence 12, Appl	770	4	40.0	192	2	US-08-290-665A-53	Sequence 53, Appl
698	4	40.0	142	3	US-08-127-721A-14	Sequence 14, Appl	771	4	40.0	192	2	US-08-290-665A-58	Sequence 58, Appl
699	4	40.0	142	3	US-08-127-721A-16	Sequence 16, Appl	772	4	40.0	192	2	US-08-290-665A-85	Sequence 85, Appl
700	4	40.0	142	3	US-08-127-721A-18	Sequence 18, Appl	773	4	40.0	192	5	PCT-US94-05795-3	Sequence 3, Appl1
701	4	40.0	142	3	US-08-485-246A-12	Sequence 12, Appl	774	4	40.0	192	5	PCT-US95-10398-53	Sequence 53, Appl
702	4	40.0	142	3	US-08-485-246A-14	Sequence 14, Appl	775	4	40.0	192	5	PCT-US95-10398-58	Sequence 58, Appl
703	4	40.0	142	3	US-08-485-246A-16	Sequence 16, Appl	776	4	40.0	192	5	PCT-US95-10398-85	Sequence 85, Appl
704	4	40.0	142	3	US-08-485-246A-18	Sequence 18, Appl	777	4	40.0	194	2	US-08-621-803-255	Sequence 255, App
705	4	40.0	142	5	PCT-US95-01219-17	Sequence 17, Appl	778	4	40.0	194	3	US-09-032-372-13	Sequence 13, Appl
706	4	40.0	144	3	US-08-545-809A-100	Sequence 100, App	779	4	40.0	194	3	US-09-078-317-12	Sequence 12, Appl
707	4	40.0	146	1	US-08-276-852-155	Sequence 155, App	780	4	40.0	195	2	US-08-621-803-263	Sequence 263, App
708	4	40.0	146	1	US-08-899-575-155	Sequence 155, App	781	4	40.0	196	4	US-09-383-586-11	Sequence 11, Appl
709	4	40.0	146	1	US-08-899-575-155	Sequence 155, App	782	4	40.0	196	1	US-08-097-827-2	Sequence 2, Appl1
710	4	40.0	146	5	PCT-US95-08743-155	Sequence 155, App	783	4	40.0	198	1	US-08-494-574-2	Sequence 2, Appl1
711	4	40.0	147	1	US-08-217-918-4	Sequence 4, Appl1	784	4	40.0	205	2	US-08-829-110-6	Sequence 6, Appl1
712	4	40.0	147	1	US-08-359-372-8	Sequence 8, Appl1	785	4	40.0	205	2	US-08-748-483-5	Sequence 5, Appl1
713	4	40.0	147	1	US-08-468-671-8	Sequence 8, Appl1	786	4	40.0	207	2	US-09-022-940-1	Sequence 1, Appl1
714	4	40.0	147	2	US-08-824-405-12	Sequence 12, Appl	787	4	40.0	207	3	US-08-996-338-23	Sequence 23, Appl
715	4	40.0	148	1	US-08-166-195A-51	Sequence 51, Appl	788	4	40.0	207	3	US-09-216-386-1	Sequence 1, Appl1
716	4	40.0	148	2	US-08-436-772-51	Sequence 51, Appl	789	4	40.0	208	2	US-08-531-525-15	Sequence 15, Appl
717	4	40.0	148	2	US-08-436-883B-51	Sequence 51, Appl	790	4	40.0	208	2	US-08-531-525-17	Sequence 17, Appl
718	4	40.0	151	2	US-08-722-050-8	Sequence 8, Appl1	791	4	40.0	208	2	US-08-844-120-1	Sequence 1, Appl1
719	4	40.0	156	1	US-07-766-682A-3	Sequence 3, Appl1	792	4	40.0	208	2	US-08-718-270A-15	Sequence 15, Appl
720	4	40.0	156	4	US-08-928-941D-4	Sequence 4, Appl1	793	4	40.0	208	2	US-08-718-270A-17	Sequence 17, Appl
721	4	40.0	156	4	US-08-928-941D-36	Sequence 36, Appl1	794	4	40.0	208	4	US-08-097-869-6	Sequence 6, Appl1
722	4	40.0	161	2	US-08-621-803-249	Sequence 249, App	795	4	40.0	210	1	US-07-667-276A-2	Sequence 2, Appl1
723	4	40.0	161	3	US-09-045-764A-4	Sequence 4, Appl1	796	4	40.0	211	4	US-08-097-869-8	Sequence 8, Appl1
724	4	40.0	161	4	US-09-271-970-2	Sequence 2, Appl1	797	4	40.0	212	2	US-08-531-525-18	Sequence 18, Appl
725	4	40.0	162	3	US-08-983-045-2	Sequence 2, Appl1	798	4	40.0	212	2	US-08-718-270A-18	Sequence 18, Appl
726	4	40.0	164	1	US-08-033-857A-8	Sequence 8, Appl1	799	4	40.0	212	4	US-08-973-462-25	Sequence 25, Appl
727	4	40.0	164	4	US-08-374-983A-8	Sequence 8, Appl1	800	4	40.0	215	2	US-08-912-129A-58	Sequence 58, Appl
728	4	40.0	164	2	US-08-824-405-2	Sequence 2, Appl1	801	4	40.0	216	2	US-08-929-418-2	Sequence 2, Appl1
729	4	40.0	166	2	US-08-483-695-3	Sequence 3, Appl1	802	4	40.0	219	1	US-08-902-516-2	Sequence 2, Appl1
730	4	40.0	166	2	US-07-965-285-3	Sequence 3, Appl1	803	4	40.0	221	1	US-08-032-848C-13	Sequence 13, Appl
731	4	40.0	166	2	US-08-487-231-3	Sequence 3, Appl1	804	4	40.0	222	1	US-08-458-516-22	Sequence 22, Appl
732	4	40.0	166	2	US-09-201-912-3	Sequence 3, Appl1	805	4	40.0	225	2	US-09-074-512-3	Sequence 3, Appl1
733	4	40.0	166	4	US-09-401-912-3	Sequence 3, Appl1	806	4	40.0	226	1	US-07-929-198-2	Sequence 2, Appl1
734	4	40.0	169	1	US-07-841-646-9	Sequence 9, Appl1	807	4	40.0	226	1	US-07-929-198-4	Sequence 4, Appl1
735	4	40.0	169	1	US-08-147-023-9	Sequence 9, Appl1	808	4	40.0	226	1	US-08-557-146-15	Sequence 15, Appl
736	4	40.0	169	1	US-08-447-570-9	Sequence 9, Appl1	809	4	40.0	226	2	US-08-557-146-15	Sequence 15, Appl
737	4	40.0	169	2	US-08-449-700-9	Sequence 9, Appl1	810	4	40.0	226	2	US-09-154-344-15	Sequence 15, Appl
738	4	40.0	169	2	US-08-449-699A-9	Sequence 9, Appl1	811	4	40.0	226	4	US-08-944-483-43	Sequence 43, Appl
739	4	40.0	173	3	US-09-188-579-87	Sequence 87, Appl	812	4	40.0	227	4	US-08-944-483-40	Sequence 40, Appl
740	4	40.0	173	4	US-09-315-444-87	Sequence 87, Appl	813	4	40.0	229	4	US-08-751-359-22	Sequence 22, Appl
741	4	40.0	174	2	US-08-683-262B-45	Sequence 45, Appl	814	4	40.0	230	1	US-08-118-469A-3	Sequence 3, Appl1
742	4	40.0	174	2	US-08-683-262B-48	Sequence 48, Appl	815	4	40.0	230	1	US-08-844-120-3	Sequence 3, Appl1
743	4	40.0	176	3	US-09-130-663-29	Sequence 29, Appl	816	4	40.0	230	2	US-08-844-120-3	Sequence 3, Appl1
744	4	40.0	176	3	US-08-392-794A-8	Sequence 8, Appl1	817	4	40.0	230	2	US-09-022-940-3	Sequence 3, Appl1
745	4	40.0	176	3	US-09-432-335-29	Sequence 29, Appl	818	4	40.0	230	2	US-09-022-940-5	Sequence 5, Appl1
746	4	40.0	178	3	US-09-081-180-2	Sequence 2, Appl1	819	4	40.0	230	3	US-09-216-001-3	Sequence 3, Appl1
747	4	40.0	178	3	US-09-040-786-2	Sequence 2, Appl1	820	4	40.0	230	3	US-09-216-386-3	Sequence 3, Appl1
748	4	40.0	178	3	US-08-705-875A-8	Sequence 8, Appl1	821	4	40.0	230	3	US-09-216-386-5	Sequence 5, Appl1
749	4	40.0	178	3	US-09-122-443-9	Sequence 9, Appl1	822	4	40.0	230	4	US-08-878-862-3	Sequence 3, Appl1
750	4	40.0	178	4	US-09-147-928-4	Sequence 4, Appl1	823	4	40.0	232	1	US-08-278-091-8	Sequence 8, Appl1
751	4	40.0	178	4	US-09-220-731-23	Sequence 23, Appl	824	4	40.0	232	1	US-08-483-859-8	Sequence 8, Appl1
752	4	40.0	178	4	US-09-220-731-24	Sequence 24, Appl	825	4	40.0	232	1	US-08-472-173-8	Sequence 8, Appl1
753	4	40.0	178	4	US-09-271-970-8	Sequence 8, Appl1	826	4	40.0	232	2	US-08-487-167-8	Sequence 8, Appl1
754	4	40.0	179	2	US-08-621-803-257	Sequence 257, App	827	4	40.0	232	2	US-08-482-816-8	Sequence 8, Appl1
755	4	40.0	180	3	US-08-621-803-261	Sequence 261, App	828	4	40.0	232	2	US-08-296-149-8	Sequence 8, Appl1
756	4	40.0	180	3	US-09-038-909-4	Sequence 4, Appl1	829	4	40.0	232	2	US-08-801-499-8	Sequence 8, Appl1
757	4	40.0	181	4	US-08-961-083-114	Sequence 114, App	830	4	40.0	232	2	US-08-978-404B-45	Sequence 45, Appl

539	4	40.0	124	1	US-08-899-575-130	Sequence 130, App	612	4	40.0	136	5	PCT-US93-11611-11	Sequence 11, Appl
540	4	40.0	124	1	US-08-899-575-131	Sequence 131, App	613	4	40.0	137	1	US-08-477-877B-96	Sequence 96, Appl
541	4	40.0	124	1	US-08-899-575-132	Sequence 132, App	614	4	40.0	137	2	US-08-472-281A-96	Sequence 96, Appl
542	4	40.0	124	1	US-08-899-575-66	Sequence 66, Appl	615	4	40.0	137	2	US-08-477-989B-96	Sequence 96, Appl
543	4	40.0	124	1	US-08-899-575-67	Sequence 67, Appl	616	4	40.0	137	3	US-08-513-968-38	Sequence 38, Appl
544	4	40.0	124	1	US-08-899-575-68	Sequence 68, Appl	617	4	40.0	138	3	US-08-513-968-38	Sequence 38, Appl
545	4	40.0	124	1	US-08-899-575-125	Sequence 125, App	618	4	40.0	139	1	US-07-718-274A-10	Sequence 2, Appl
546	4	40.0	124	1	US-08-899-575-126	Sequence 126, App	619	4	40.0	139	1	US-08-163-877-12	Sequence 10, Appl
547	4	40.0	124	1	US-08-899-575-127	Sequence 127, App	620	4	40.0	139	1	US-08-149-106-2	Sequence 2, Appl
548	4	40.0	124	1	US-08-899-575-130	Sequence 130, App	621	4	40.0	139	1	US-08-298-021-2	Sequence 2, Appl
549	4	40.0	124	1	US-08-899-575-131	Sequence 131, App	622	4	40.0	139	1	US-08-278-729A-5	Sequence 5, Appl
550	4	40.0	124	1	US-08-899-575-132	Sequence 132, App	623	4	40.0	139	1	US-08-278-729A-5	Sequence 5, Appl
551	4	40.0	124	5	PCT-US95-08743-66	Sequence 66, Appl	624	4	40.0	139	1	US-08-155-343A-5	Sequence 5, Appl
552	4	40.0	124	5	PCT-US95-08743-67	Sequence 67, Appl	625	4	40.0	139	1	US-08-155-343A-6	Sequence 6, Appl
553	4	40.0	124	5	PCT-US95-08743-68	Sequence 68, Appl	626	4	40.0	139	1	US-08-406-672-6	Sequence 6, Appl
554	4	40.0	124	5	PCT-US95-08743-125	Sequence 125, App	627	4	40.0	139	1	US-08-406-672-6	Sequence 6, Appl
555	4	40.0	124	5	PCT-US95-08743-126	Sequence 126, App	628	4	40.0	139	1	US-08-643-553A-5	Sequence 5, Appl
556	4	40.0	124	5	PCT-US95-08743-127	Sequence 127, App	629	4	40.0	139	1	US-08-643-553A-6	Sequence 6, Appl
557	4	40.0	124	5	PCT-US95-08743-130	Sequence 130, App	630	4	40.0	139	1	US-08-643-553A-5	Sequence 5, Appl
558	4	40.0	124	5	PCT-US95-08743-131	Sequence 131, App	631	4	40.0	139	1	US-08-643-763A-6	Sequence 6, Appl
559	4	40.0	124	5	PCT-US95-08743-132	Sequence 132, App	632	4	40.0	139	1	US-08-462-623-5	Sequence 5, Appl
560	4	40.0	125	1	US-08-276-852-124	Sequence 124, App	633	4	40.0	139	1	US-08-462-623-6	Sequence 6, Appl
561	4	40.0	125	1	US-08-276-852-128	Sequence 128, App	634	4	40.0	139	1	US-08-451-953A-5	Sequence 5, Appl
562	4	40.0	125	1	US-08-276-852-129	Sequence 129, App	635	4	40.0	139	1	US-08-451-953A-6	Sequence 6, Appl
563	4	40.0	125	1	US-08-899-575-124	Sequence 124, App	636	4	40.0	139	1	US-08-360-914B-10	Sequence 10, Appl
564	4	40.0	125	1	US-08-899-575-128	Sequence 128, App	637	4	40.0	139	1	US-08-253-877C-19	Sequence 19, Appl
565	4	40.0	125	1	US-08-899-575-129	Sequence 129, App	638	4	40.0	139	1	US-08-202-047-7	Sequence 7, Appl
566	4	40.0	125	1	US-08-899-575-124	Sequence 124, App	639	4	40.0	139	1	US-08-202-047-9	Sequence 9, Appl
567	4	40.0	125	1	US-08-899-575-128	Sequence 128, App	640	4	40.0	139	1	US-08-202-047-11	Sequence 11, Appl
568	4	40.0	125	1	US-08-899-575-129	Sequence 129, App	641	4	40.0	139	1	US-08-741-589A-10	Sequence 10, Appl
569	4	40.0	125	2	US-08-665-202-56	Sequence 56, Appl	642	4	40.0	139	2	US-08-656-586-8	Sequence 8, Appl
570	4	40.0	125	2	US-08-665-202-58	Sequence 58, Appl	643	4	40.0	139	2	US-08-445-468A-5	Sequence 5, Appl
571	4	40.0	125	2	US-08-665-202-59	Sequence 59, Appl	644	4	40.0	139	2	US-08-445-468A-6	Sequence 6, Appl
572	4	40.0	125	4	US-09-199-149-3	Sequence 3, Appl	645	4	40.0	139	2	US-08-452-164A-19	Sequence 19, Appl
573	4	40.0	125	5	PCT-US95-08743-124	Sequence 124, App	646	4	40.0	139	2	US-08-461-357A-5	Sequence 5, Appl
574	4	40.0	125	5	PCT-US95-08743-128	Sequence 128, App	647	4	40.0	139	2	US-08-461-357A-6	Sequence 6, Appl
575	4	40.0	125	5	PCT-US95-08743-129	Sequence 129, App	648	4	40.0	139	2	US-08-912-088-5	Sequence 5, Appl
576	4	40.0	126	1	US-08-276-852-123	Sequence 123, App	649	4	40.0	139	2	US-08-912-088-6	Sequence 6, Appl
577	4	40.0	126	1	US-08-899-575-123	Sequence 123, App	650	4	40.0	139	3	US-08-603-082-18	Sequence 18, Appl
578	4	40.0	126	1	US-08-899-575-123	Sequence 123, App	651	4	40.0	139	3	US-08-278-720A-5	Sequence 5, Appl
579	4	40.0	126	5	PCT-US95-08743-123	Sequence 123, App	652	4	40.0	139	3	US-08-278-720A-6	Sequence 6, Appl
580	4	40.0	128	1	US-08-276-852-56	Sequence 56, Appl	653	4	40.0	139	3	US-08-964-650-9	Sequence 9, Appl
581	4	40.0	128	1	US-08-276-852-58	Sequence 58, Appl	654	4	40.0	139	3	US-08-964-650-9	Sequence 9, Appl
582	4	40.0	128	1	US-08-478-039-63	Sequence 63, Appl	655	4	40.0	139	3	US-08-964-650-11	Sequence 11, Appl
583	4	40.0	128	1	US-08-476-349A-63	Sequence 63, Appl	656	4	40.0	139	3	US-08-445-467-5	Sequence 5, Appl
584	4	40.0	128	1	US-08-899-575-56	Sequence 56, Appl	657	4	40.0	139	3	US-08-445-467-6	Sequence 6, Appl
585	4	40.0	128	1	US-08-899-575-58	Sequence 58, Appl	658	4	40.0	139	3	US-08-443-676-1	Sequence 1, Appl
586	4	40.0	128	1	US-08-202-047-22	Sequence 22, Appl	659	4	40.0	139	3	US-08-480-515A-5	Sequence 5, Appl
587	4	40.0	128	1	US-08-899-575-56	Sequence 56, Appl	660	4	40.0	139	3	US-08-480-515A-6	Sequence 6, Appl
588	4	40.0	128	3	US-08-899-575-58	Sequence 58, Appl	661	4	40.0	139	3	US-08-933-983-21	Sequence 21, Appl
589	4	40.0	128	3	US-08-964-650-22	Sequence 22, Appl	662	4	40.0	139	4	US-08-414-033A-5	Sequence 5, Appl
590	4	40.0	128	4	US-09-199-149-12	Sequence 12, Appl	663	4	40.0	139	4	US-08-414-033A-6	Sequence 6, Appl
591	4	40.0	128	5	PCT-US95-08743-56	Sequence 56, Appl	664	4	40.0	139	4	US-08-271-556A-3	Sequence 3, Appl
592	4	40.0	128	5	PCT-US95-08743-58	Sequence 58, Appl	665	4	40.0	139	4	US-08-271-556A-4	Sequence 4, Appl
593	4	40.0	129	1	US-08-313-075A-52	Sequence 52, Appl	666	4	40.0	139	4	US-08-905-223-49	Sequence 49, App
594	4	40.0	129	2	US-08-561-521-45	Sequence 45, Appl	667	4	40.0	139	5	PCT-US92-01968-5	Sequence 5, Appl
595	4	40.0	129	5	PCT-US95-01219-45	Sequence 45, Appl	668	4	40.0	139	5	PCT-US92-01968-6	Sequence 6, Appl
596	4	40.0	133	4	US-08-718-333A-8	Sequence 8, Appl	669	4	40.0	139	5	PCT-US93-07190-5	Sequence 5, Appl
597	4	40.0	133	4	US-07-634-278-19	Sequence 19, Appl	670	4	40.0	139	5	PCT-US93-07190-6	Sequence 6, Appl
598	4	40.0	135	1	US-08-477-728-19	Sequence 19, Appl	671	4	40.0	139	5	PCT-US93-07231-5	Sequence 5, Appl
599	4	40.0	135	1	US-08-474-040-19	Sequence 19, Appl	672	4	40.0	139	5	PCT-US93-07231-6	Sequence 6, Appl
600	4	40.0	135	1	US-08-487-200-19	Sequence 19, App	673	4	40.0	139	5	PCT-US93-08742-5	Sequence 5, Appl
601	4	40.0	135	1	US-08-137-117D-100	Sequence 100, App	674	4	40.0	139	5	PCT-US93-08742-6	Sequence 6, Appl
602	4	40.0	135	1	US-08-137-117D-102	Sequence 102, App	675	4	40.0	139	5	PCT-US93-08808-5	Sequence 5, Appl
603	4	40.0	135	2	US-08-137-117D-112	Sequence 112, App	676	4	40.0	139	5	PCT-US93-08808-6	Sequence 6, Appl
604	4	40.0	135	2	US-08-436-717-100	Sequence 100, App	677	4	40.0	139	5	PCT-US93-08885-5	Sequence 5, Appl
605	4	40.0	135	2	US-08-436-717-102	Sequence 102, App	678	4	40.0	139	5	PCT-US93-08885-6	Sequence 6, Appl
606	4	40.0	135	2	US-08-303-569B-31	Sequence 31, App	679	4	40.0	139	5	PCT-US94-13181-10	Sequence 10, Appl
607	4	40.0	135	2	US-08-621-751A-16	Sequence 16, App	680	4	40.0	140	1	US-07-946-421-28	Sequence 28, Appl
608	4	40.0	135	2	US-08-484-537-19	Sequence 19, App	681	4	40.0	140	3	US-08-836-561-63	Sequence 63, Appl
609	4	40.0	135	2	US-08-649-100-41	Sequence 41, App	682	4	40.0	140	3	US-08-836-561-78	Sequence 78, Appl
610	4	40.0	136	3	US-08-646-265A-99	Sequence 99, App	683	4	40.0	140	3	US-08-836-561-83	Sequence 83, Appl
611	4	40.0	136	4	US-08-646-265A-99	Sequence 99, App	684	4	40.0	140	3	US-08-836-561-83	Sequence 83, Appl

393	4	40.0	117	4	US-09-042-353-232	Sequence 232, App	466	4	40.0	120	3	US-08-397-411-10	Sequence 10, App1
394	4	40.0	117	5	PCT-US93-11611-6	Sequence 7, App1	467	4	40.0	120	5	PCT-US93-10555-12	Sequence 12, App1
395	4	40.0	117	5	PCT-US93-11611-7	Sequence 6, App1	468	4	40.0	120	5	PCT-US93-10555-13	Sequence 13, App1
396	4	40.0	117	5	PCT-US95-10053-14	Sequence 14, App1	469	4	40.0	121	1	US-07-634-278-53	Sequence 53, App1
397	4	40.0	117	5	PCT-US96-09448-17	Sequence 17, App1	470	4	40.0	121	1	US-08-477-728-53	Sequence 53, App1
398	4	40.0	118	1	US-08-500-386A-64	Sequence 64, App1	471	4	40.0	121	1	US-08-264-093-3	Sequence 3, App1
399	4	40.0	118	1	US-08-477-877B-93	Sequence 93, App1	472	4	40.0	121	1	US-08-478-039-64	Sequence 64, App1
400	4	40.0	118	1	US-08-491-845-14	Sequence 14, App1	473	4	40.0	121	1	US-08-474-040-53	Sequence 53, App1
401	4	40.0	118	1	US-08-481-377-12	Sequence 12, App1	474	4	40.0	121	1	US-08-487-200-64	Sequence 64, App1
402	4	40.0	118	2	US-08-472-281A-93	Sequence 93, App1	475	4	40.0	121	1	US-08-476-349A-53	Sequence 53, App1
403	4	40.0	118	2	US-08-491-835-10	Sequence 10, App1	476	4	40.0	121	1	US-08-202-047-23	Sequence 23, App1
404	4	40.0	118	2	US-08-428-257A-78	Sequence 78, App1	477	4	40.0	121	1	US-08-420-235B-23	Sequence 23, App1
405	4	40.0	118	2	US-08-232-081B-8	Sequence 8, App1	478	4	40.0	121	2	US-08-232-081B-41	Sequence 41, App1
406	4	40.0	118	2	US-08-477-989B-93	Sequence 93, App1	479	4	40.0	121	2	US-08-428-197-14	Sequence 14, App1
407	4	40.0	118	3	US-08-392-794A-4	Sequence 4, App1	480	4	40.0	121	3	US-08-964-690-23	Sequence 23, App1
408	4	40.0	118	3	US-09-153-733A-12	Sequence 12, App1	481	4	40.0	121	4	US-08-983-607-20	Sequence 20, App1
409	4	40.0	118	3	US-08-946-092A-10	Sequence 10, App1	482	4	40.0	121	4	US-08-793-624-23	Sequence 23, App1
410	4	40.0	118	3	US-08-931-645-64	Sequence 64, App1	483	4	40.0	121	4	US-08-484-537-82	Sequence 82, App1
411	4	40.0	118	3	US-08-513-96B-80	Sequence 80, App1	484	4	40.0	121	4	US-08-579-378A-8	Sequence 8, App1
412	4	40.0	118	4	US-08-983-607-24	Sequence 24, App1	485	4	40.0	121	5	PCT-US93-10555-14	Sequence 14, App1
413	4	40.0	118	4	US-09-199-149-5	Sequence 5, App1	486	4	40.0	121	5	PCT-US93-11612-8	Sequence 8, App1
414	4	40.0	118	4	US-09-172-062-10	Sequence 10, App1	487	4	40.0	121	5	PCT-US95-10194-23	Sequence 23, App1
415	4	40.0	118	4	US-07-687-264-60	Sequence 60, App1	488	4	40.0	121	5	PCT-US95-1365B-4	Sequence 4, App1
416	4	40.0	118	5	PCT-US94-00666-12	Sequence 12, App1	489	4	40.0	122	1	US-07-634-278-88	Sequence 88, App1
417	4	40.0	118	5	PCT-US94-00685-10	Sequence 10, App1	490	4	40.0	122	1	US-08-477-728-88	Sequence 88, App1
418	4	40.0	118	5	PCT-US95-11235-64	Sequence 64, App1	491	4	40.0	122	1	US-08-474-040-88	Sequence 88, App1
419	4	40.0	119	1	US-07-634-278-65	Sequence 65, App1	492	4	40.0	122	1	US-08-487-200-88	Sequence 88, App1
420	4	40.0	119	1	US-07-634-278-89	Sequence 89, App1	493	4	40.0	122	4	US-08-484-537-88	Sequence 88, App1
421	4	40.0	119	1	US-08-477-728-65	Sequence 65, App1	494	4	40.0	122	5	PCT-US95-00067-2	Sequence 2, App1
422	4	40.0	119	1	US-08-477-728-89	Sequence 89, App1	495	4	40.0	122	6	5304466-3	Patent No. 5304466
423	4	40.0	119	1	US-08-300-386A-65	Sequence 65, App1	496	4	40.0	123	1	US-08-497-312-17	Sequence 17, App1
424	4	40.0	119	1	US-08-478-039-65	Sequence 65, App1	497	4	40.0	123	1	US-08-477-877B-94	Sequence 94, App1
425	4	40.0	119	1	US-08-474-040-65	Sequence 65, App1	498	4	40.0	123	1	US-08-482-882-53	Sequence 53, App1
426	4	40.0	119	1	US-08-474-040-89	Sequence 89, App1	499	4	40.0	123	1	US-08-482-882-86	Sequence 86, App1
427	4	40.0	119	1	US-08-487-200-65	Sequence 65, App1	500	4	40.0	123	2	US-08-483-389-53	Sequence 53, App1
428	4	40.0	119	1	US-08-487-200-89	Sequence 89, App1	501	4	40.0	123	2	US-08-483-389-86	Sequence 86, App1
429	4	40.0	119	1	US-08-488-113B-160	Sequence 160, App1	502	4	40.0	123	2	US-08-472-881A-94	Sequence 94, App1
430	4	40.0	119	1	US-08-476-349A-65	Sequence 65, App1	503	4	40.0	123	2	US-08-487-113B-53	Sequence 53, App1
431	4	40.0	119	1	US-08-477-484B-160	Sequence 160, App1	504	4	40.0	123	2	US-08-487-113B-86	Sequence 86, App1
432	4	40.0	119	1	US-08-107-669B-24	Sequence 24, App1	505	4	40.0	123	2	US-08-561-521-11	Sequence 11, App1
433	4	40.0	119	1	US-08-472-788A-24	Sequence 24, App1	506	4	40.0	123	2	US-08-473-503-53	Sequence 53, App1
434	4	40.0	119	1	US-08-581-529B-12	Sequence 12, App1	507	4	40.0	123	2	US-08-473-503-86	Sequence 86, App1
435	4	40.0	119	1	US-08-491-845-10	Sequence 10, App1	508	4	40.0	123	2	US-08-652-816A-6	Sequence 6, App1
436	4	40.0	119	1	US-08-458-516-10	Sequence 10, App1	509	4	40.0	123	2	US-08-652-816A-7	Sequence 7, App1
437	4	40.0	119	2	US-08-455-559-18	Sequence 18, App1	510	4	40.0	123	2	US-08-652-816A-8	Sequence 8, App1
438	4	40.0	119	2	US-08-455-559-24	Sequence 24, App1	511	4	40.0	123	2	US-08-652-816A-9	Sequence 9, App1
439	4	40.0	119	2	US-08-525-596B-22	Sequence 22, App1	512	4	40.0	123	2	US-08-652-816A-9	Sequence 9, App1
440	4	40.0	119	2	US-08-546-360-160	Sequence 160, App1	513	4	40.0	123	2	US-08-483-932-53	Sequence 53, App1
441	4	40.0	119	2	US-08-561-521-10	Sequence 10, App1	514	4	40.0	123	2	US-08-483-932-86	Sequence 86, App1
442	4	40.0	119	2	US-08-561-521-12	Sequence 12, App1	515	4	40.0	123	2	US-08-477-989B-94	Sequence 94, App1
443	4	40.0	119	2	US-08-561-521-13	Sequence 13, App1	516	4	40.0	123	2	US-08-720-420A-53	Sequence 53, App1
444	4	40.0	119	2	US-08-082-842A-24	Sequence 24, App1	517	4	40.0	123	2	US-08-720-420A-86	Sequence 86, App1
445	4	40.0	119	2	US-08-581-528A-12	Sequence 12, App1	518	4	40.0	123	3	US-08-714-017-53	Sequence 53, App1
446	4	40.0	119	3	US-09-097-616-12	Sequence 12, App1	519	4	40.0	123	3	US-08-714-017-86	Sequence 86, App1
447	4	40.0	119	3	US-09-177-860A-22	Sequence 22, App1	520	4	40.0	123	3	US-08-475-680-53	Sequence 53, App1
448	4	40.0	119	3	US-08-931-645-65	Sequence 65, App1	521	4	40.0	123	3	US-08-475-680-86	Sequence 86, App1
449	4	40.0	119	4	US-08-983-607-50	Sequence 50, App1	522	4	40.0	123	4	US-09-102-528-22	Sequence 22, App1
450	4	40.0	119	4	US-08-839-765-160	Sequence 160, App1	523	4	40.0	123	5	PCT-US95-01219-11	Sequence 11, App1
451	4	40.0	119	4	US-09-136-389-160	Sequence 160, App1	524	4	40.0	124	1	US-08-276-852-66	Sequence 66, App1
452	4	40.0	119	4	US-08-484-537-65	Sequence 65, App1	525	4	40.0	124	1	US-08-276-852-67	Sequence 67, App1
453	4	40.0	119	4	US-08-484-537-89	Sequence 89, App1	526	4	40.0	124	1	US-08-276-852-68	Sequence 68, App1
454	4	40.0	119	4	US-08-624-635-13	Sequence 13, App1	527	4	40.0	124	1	US-08-276-852-125	Sequence 125, App1
455	4	40.0	119	4	US-09-145-060-18	Sequence 18, App1	528	4	40.0	124	1	US-08-276-852-126	Sequence 126, App1
456	4	40.0	119	5	PCT-US94-00667-18	Sequence 18, App1	529	4	40.0	124	1	US-08-276-852-127	Sequence 127, App1
457	4	40.0	119	5	PCT-US94-07762-12	Sequence 12, App1	530	4	40.0	124	1	US-08-276-852-130	Sequence 130, App1
458	4	40.0	119	5	PCT-US94-07799-12	Sequence 12, App1	531	4	40.0	124	1	US-08-276-852-131	Sequence 131, App1
459	4	40.0	119	5	PCT-US95-01219-10	Sequence 10, App1	532	4	40.0	124	1	US-08-276-852-132	Sequence 132, App1
460	4	40.0	119	5	PCT-US95-01219-12	Sequence 12, App1	533	4	40.0	124	1	US-08-899-575-66	Sequence 66, App1
461	4	40.0	119	5	PCT-US95-01219-13	Sequence 13, App1	534	4	40.0	124	1	US-08-899-575-67	Sequence 67, App1
462	4	40.0	119	5	PCT-US95-11235-65	Sequence 65, App1	535	4	40.0	124	1	US-08-899-575-68	Sequence 68, App1
463	4	40.0	120	2	US-08-652-816A-19	Sequence 19, App1	536	4	40.0	124	1	US-08-899-575-125	Sequence 125, App1
464	4	40.0	120	2	US-08-428-197-12	Sequence 12, App1	537	4	40.0	124	1	US-08-899-575-126	Sequence 126, App1
465	4	40.0	120	2	US-08-428-197-13	Sequence 13, App1	538	4	40.0	124	1	US-08-899-575-127	Sequence 127, App1

247	4	40.0	94	1	US-08-167-035-23	Sequence 23, Appl	320	4	40.0	116	1	US-08-487-200-57	Sequence 57, Appl
248	4	40.0	94	1	US-08-208-887A-23	Sequence 23, Appl	321	4	40.0	116	1	US-08-487-200-73	Sequence 73, Appl
249	4	40.0	94	2	US-08-539-005-23	Sequence 23, Appl	322	4	40.0	116	1	US-08-488-113B-169	Sequence 169, App
250	4	40.0	97	2	US-08-290-592E-16	Sequence 16, Appl	323	4	40.0	116	1	US-08-477-444B-169	Sequence 169, App
251	4	40.0	97	5	PCT-US95-10053-13	Sequence 13, Appl	324	4	40.0	116	1	US-08-107-669D-55	Sequence 55, Appl
252	4	40.0	97	5	PCT-US96-09448-16	Sequence 16, Appl	325	4	40.0	116	1	US-08-472-788A-86	Sequence 86, Appl
253	4	40.0	98	1	US-08-211-202-130	Sequence 130, App	326	4	40.0	116	1	US-08-690-102A-8	Sequence 8, Appl
254	4	40.0	98	1	US-08-211-202-140	Sequence 140, App	327	4	40.0	116	1	US-08-690-102A-9	Sequence 9, Appl
255	4	40.0	98	1	US-08-478-039-81	Sequence 81, Appl	328	4	40.0	116	2	US-08-477-531B-55	Sequence 55, Appl
256	4	40.0	98	1	US-08-476-349A-81	Sequence 81, Appl	329	4	40.0	116	2	US-08-646-380-169	Sequence 169, App
257	4	40.0	98	2	US-08-428-197-17	Sequence 17, Appl	330	4	40.0	116	2	US-08-561-521-41	Sequence 41, Appl
258	4	40.0	98	2	US-08-665-202-33	Sequence 33, Appl	331	4	40.0	116	2	US-08-082-842A-86	Sequence 86, Appl
259	4	40.0	98	5	PCT-US93-10555-17	Sequence 17, Appl	332	4	40.0	116	3	US-08-934-841-1	Sequence 1, Appl
260	4	40.0	100	3	US-09-034-916-13	Sequence 13, Appl	333	4	40.0	116	3	US-08-545-809A-104	Sequence 104, App
261	4	40.0	100	4	US-08-973-462-11	Sequence 11, Appl	334	4	40.0	116	3	US-09-393-809A-1	Sequence 1, Appl
262	4	40.0	102	1	US-07-834-539A-55	Sequence 55, Appl	335	4	40.0	116	4	US-08-839-765-169	Sequence 169, App
263	4	40.0	102	1	US-08-033-131-63	Sequence 63, Appl	336	4	40.0	116	4	US-09-136-389-169	Sequence 169, App
264	4	40.0	102	1	US-08-335-583C-47	Sequence 47, Appl	337	4	40.0	116	4	US-08-484-537-6	Sequence 6, Appl
265	4	40.0	102	1	US-08-645-641-63	Sequence 63, Appl	338	4	40.0	116	4	US-08-484-537-57	Sequence 57, Appl
266	4	40.0	102	1	US-07-853-408B-63	Sequence 63, Appl	339	4	40.0	116	4	US-08-484-537-73	Sequence 73, Appl
267	4	40.0	102	2	US-08-096-762-63	Sequence 63, Appl	340	4	40.0	116	4	US-08-127-902-8	Sequence 8, Appl
268	4	40.0	102	2	US-08-800-353-55	Sequence 55, Appl	341	4	40.0	116	4	US-09-127-902-9	Sequence 9, Appl
269	4	40.0	102	2	US-08-308-865-63	Sequence 63, Appl	342	4	40.0	116	4	US-09-127-902-9	Sequence 9, Appl
270	4	40.0	102	2	US-08-288-508C-18	Sequence 18, Appl	343	4	40.0	116	4	US-09-155-107-8	Sequence 8, Appl
271	4	40.0	102	3	US-08-478-097A-16	Sequence 16, Appl	344	4	40.0	116	4	US-09-155-107-21	Sequence 21, Appl
272	4	40.0	102	3	US-08-289-222E-27	Sequence 27, Appl	345	4	40.0	116	5	PCT-US95-01219-41	Sequence 41, Appl
273	4	40.0	102	4	US-09-054-526B-27	Sequence 27, Appl	346	4	40.0	116	5	PCT-US95-09641-8	Sequence 8, Appl
274	4	40.0	102	4	US-08-931-858E-161	Sequence 161, App	347	4	40.0	116	5	PCT-US95-09641-9	Sequence 9, Appl
275	4	40.0	102	4	US-08-981-739-161	Sequence 161, App	348	4	40.0	117	1	US-07-634-278-4	Sequence 4, Appl
276	4	40.0	102	5	PCT-US92-06185-55	Sequence 55, Appl	349	4	40.0	117	1	US-07-634-278-15	Sequence 15, Appl
277	4	40.0	102	5	PCT-US92-10983-63	Sequence 63, Appl	350	4	40.0	117	1	US-07-634-278-72	Sequence 72, Appl
278	4	40.0	104	1	US-07-764-731B-8	Sequence 8, Appl	351	4	40.0	117	1	US-07-634-278-104	Sequence 104, App
279	4	40.0	104	3	US-08-894-173-71	Sequence 71, Appl	352	4	40.0	117	1	US-07-634-278-105	Sequence 105, App
280	4	40.0	104	4	US-09-398-193-71	Sequence 71, Appl	353	4	40.0	117	1	US-08-477-728-4	Sequence 4, Appl
281	4	40.0	105	1	US-08-422-101-9	Sequence 9, Appl	354	4	40.0	117	1	US-08-477-728-15	Sequence 15, Appl
282	4	40.0	105	1	US-08-422-091-9	Sequence 9, Appl	355	4	40.0	117	1	US-08-477-728-72	Sequence 72, Appl
283	4	40.0	105	2	US-08-422-092-9	Sequence 9, Appl	356	4	40.0	117	1	US-08-477-728-104	Sequence 104, App
284	4	40.0	105	2	US-08-788-800-6	Sequence 6, Appl	357	4	40.0	117	1	US-08-477-728-105	Sequence 105, App
285	4	40.0	105	3	US-08-423-093-9	Sequence 9, Appl	358	4	40.0	117	1	US-08-474-040-4	Sequence 4, Appl
286	4	40.0	105	3	US-08-422-112-9	Sequence 9, Appl	359	4	40.0	117	1	US-08-474-040-15	Sequence 15, Appl
287	4	40.0	109	2	US-08-761-277A-51	Sequence 51, Appl	360	4	40.0	117	1	US-08-474-040-12	Sequence 12, Appl
288	4	40.0	109	3	US-08-894-173-88	Sequence 88, Appl	361	4	40.0	117	1	US-08-474-040-104	Sequence 104, App
289	4	40.0	109	4	US-09-398-193-88	Sequence 88, Appl	362	4	40.0	117	1	US-08-474-040-105	Sequence 105, App
290	4	40.0	110	1	US-08-466-886-21	Sequence 21, Appl	363	4	40.0	117	1	US-08-487-200-4	Sequence 4, Appl
291	4	40.0	110	1	US-08-466-886-23	Sequence 23, Appl	364	4	40.0	117	1	US-08-487-200-15	Sequence 15, Appl
292	4	40.0	110	1	US-08-466-886-25	Sequence 25, Appl	365	4	40.0	117	1	US-08-487-200-72	Sequence 72, Appl
293	4	40.0	110	1	US-08-961-083-102	Sequence 102, App	366	4	40.0	117	1	US-08-487-200-104	Sequence 104, App
294	4	40.0	110	4	US-08-469-617-21	Sequence 21, Appl	367	4	40.0	117	1	US-08-487-200-105	Sequence 105, App
295	4	40.0	110	4	US-08-469-617-23	Sequence 23, Appl	368	4	40.0	117	1	US-08-488-113B-166	Sequence 166, App
296	4	40.0	110	4	US-08-469-617-25	Sequence 25, Appl	369	4	40.0	117	1	US-08-477-484B-166	Sequence 166, App
297	4	40.0	112	3	US-08-545-809A-113	Sequence 113, App	370	4	40.0	117	1	US-08-107-669D-52	Sequence 52, Appl
298	4	40.0	113	4	US-08-836-075A-70	Sequence 70, Appl	371	4	40.0	117	1	US-08-472-788A-52	Sequence 52, Appl
299	4	40.0	113	4	US-08-836-075A-72	Sequence 72, Appl	372	4	40.0	117	2	US-08-477-531B-52	Sequence 52, Appl
300	4	40.0	113	4	US-08-836-075A-74	Sequence 74, Appl	373	4	40.0	117	2	US-08-290-592E-17	Sequence 17, Appl
301	4	40.0	113	4	US-08-836-075A-78	Sequence 78, Appl	374	4	40.0	117	2	US-08-646-360-166	Sequence 166, App
302	4	40.0	114	4	US-09-188-930-161	Sequence 161, App	375	4	40.0	117	2	US-08-082-842A-52	Sequence 52, Appl
303	4	40.0	114	4	US-09-188-930-288	Sequence 288, App	376	4	40.0	117	3	US-08-545-809A-90	Sequence 90, Appl
304	4	40.0	115	2	US-07-903-029-6	Sequence 6, Appl	377	4	40.0	117	3	US-08-545-809A-91	Sequence 91, Appl
305	4	40.0	115	3	US-08-478-057A-30	Sequence 30, Appl	378	4	40.0	117	3	US-08-545-809A-96	Sequence 96, Appl
306	4	40.0	116	1	US-07-634-278-5	Sequence 5, Appl	379	4	40.0	117	3	US-08-545-809A-105	Sequence 105, App
307	4	40.0	116	1	US-07-634-278-6	Sequence 6, Appl	380	4	40.0	117	3	US-08-545-809A-110	Sequence 110, App
308	4	40.0	116	1	US-07-634-278-57	Sequence 57, Appl	381	4	40.0	117	3	US-08-545-809A-127	Sequence 127, App
309	4	40.0	116	1	US-07-634-278-73	Sequence 73, Appl	382	4	40.0	117	3	US-08-545-809A-128	Sequence 128, App
310	4	40.0	116	1	US-08-477-728-5	Sequence 5, Appl	383	4	40.0	117	3	US-08-545-809A-133	Sequence 133, App
311	4	40.0	116	1	US-08-477-728-6	Sequence 6, Appl	384	4	40.0	117	3	US-08-545-809A-139	Sequence 139, App
312	4	40.0	116	1	US-08-477-728-57	Sequence 57, Appl	385	4	40.0	117	4	US-08-839-765-166	Sequence 166, App
313	4	40.0	116	1	US-08-477-728-73	Sequence 73, Appl	386	4	40.0	117	4	US-09-136-389-166	Sequence 166, App
314	4	40.0	116	1	US-08-474-040-5	Sequence 5, Appl	387	4	40.0	117	4	US-08-484-537-4	Sequence 4, Appl
315	4	40.0	116	1	US-08-474-040-6	Sequence 6, Appl	388	4	40.0	117	4	US-08-484-537-15	Sequence 15, Appl
316	4	40.0	116	1	US-08-474-040-57	Sequence 57, Appl	389	4	40.0	117	4	US-08-484-537-72	Sequence 72, Appl
317	4	40.0	116	1	US-08-474-040-73	Sequence 73, Appl	390	4	40.0	117	4	US-08-484-537-104	Sequence 104, App
318	4	40.0	116	1	US-08-487-200-5	Sequence 5, Appl	391	4	40.0	117	4	US-08-484-537-105	Sequence 105, App
319	4	40.0	116	1	US-08-487-200-6	Sequence 6, Appl	392	4	40.0	117	4	US-08-646-265A-132	Sequence 132, App

101	4	40.0	9	5	PCT-US94-00616-22	Sequence 22, Appl	174	4	40.0	29	3	US-08-374-077C-16	Sequence 16, Appl
102	4	40.0	10	1	US-08-250-789A-55	Sequence 55, Appl	175	4	40.0	29	4	US-08-895-590-58	Sequence 58, Appl
103	4	40.0	11	1	US-08-336-343A-29	Sequence 29, Appl	176	4	40.0	30	1	US-07-946-421-11	Sequence 11, Appl
104	4	40.0	11	1	US-08-338-634-8	Sequence 8, Appl	177	4	40.0	30	1	US-08-477-877B-52	Sequence 52, Appl
105	4	40.0	11	2	US-08-456-670B-38	Sequence 38, Appl	178	4	40.0	30	1	US-08-137-117D-146	Sequence 146, App
106	4	40.0	11	2	US-08-705-875A-11	Sequence 11, Appl	179	4	40.0	30	1	US-08-137-117D-152	Sequence 152, App
107	4	40.0	11	4	US-08-652-877-29	Sequence 29, Appl	180	4	40.0	30	1	US-08-137-117D-157	Sequence 157, App
108	4	40.0	11	4	US-08-476-515A-29	Sequence 29, Appl	181	4	40.0	30	2	US-08-472-281A-52	Sequence 52, Appl
109	4	40.0	11	4	US-09-001-984C-7	Sequence 7, Appl	182	4	40.0	30	2	US-08-436-717-145	Sequence 145, App
110	4	40.0	11	4	US-09-001-984C-23	Sequence 23, Appl	183	4	40.0	30	2	US-08-436-717-152	Sequence 152, App
111	4	40.0	11	5	US-09-001-984C-55	Sequence 55, Appl	184	4	40.0	30	2	US-08-436-717-157	Sequence 157, App
112	4	40.0	11	5	PCT-US95-16415-18	Sequence 18, Appl	185	4	40.0	30	2	US-08-560-558B-5	Sequence 5, Appl
113	4	40.0	12	4	US-08-602-999A-255	Sequence 255, App	186	4	40.0	30	2	US-08-477-989B-52	Sequence 52, Appl
114	4	40.0	13	3	US-08-705-875A-12	Sequence 12, Appl	187	4	40.0	30	2	US-08-287-537B-7	Sequence 7, Appl
115	4	40.0	13	3	US-09-001-984C-8	Sequence 8, Appl	188	4	40.0	30	2	US-08-287-537B-8	Sequence 8, Appl
116	4	40.0	13	4	US-09-001-984C-9	Sequence 9, Appl	189	4	40.0	30	3	US-08-649-100-34	Sequence 34, Appl
117	4	40.0	13	4	US-09-001-984C-20	Sequence 20, Appl	190	4	40.0	30	4	US-08-569-147-54	Sequence 54, Appl
118	4	40.0	13	4	US-09-001-984C-25	Sequence 25, Appl	191	4	40.0	30	4	US-08-646-265A-126	Sequence 126, App
119	4	40.0	14	3	US-08-705-875A-13	Sequence 13, Appl	192	4	40.0	30	5	PCT-US94-10936-7	Sequence 7, Appl
120	4	40.0	15	1	US-08-408-604A-31	Sequence 31, Appl	193	4	40.0	30	5	PCT-US94-10936-8	Sequence 8, Appl
121	4	40.0	15	2	US-08-687-956A-3	Sequence 3, Appl	194	4	40.0	31	1	US-08-488-252-34	Sequence 34, Appl
122	4	40.0	15	2	US-08-687-956A-6	Sequence 6, Appl	195	4	40.0	31	2	US-08-023-980B-30	Sequence 30, Appl
123	4	40.0	15	4	US-08-595-945-11	Sequence 11, Appl	196	4	40.0	31	2	US-08-486-953A-25	Sequence 25, Appl
124	4	40.0	15	4	US-08-602-999A-410	Sequence 410, App	197	4	40.0	31	6	5512648-2	Patent No. 5512648
125	4	40.0	15	5	PCT-US93-09626-31	Sequence 31, Appl	198	4	40.0	32	2	US-08-780-836B-2	Sequence 2, Appl
126	4	40.0	16	1	US-08-447-925-5	Sequence 5, Appl	199	4	40.0	33	1	US-08-086-428B-137	Sequence 137, App
127	4	40.0	16	1	US-08-469-615-19	Sequence 19, Appl	200	4	40.0	33	2	US-08-468-570-137	Sequence 137, App
128	4	40.0	16	1	US-08-466-763-19	Sequence 19, Appl	201	4	40.0	33	2	US-08-290-665A-241	Sequence 241, App
129	4	40.0	16	2	US-08-411-142A-19	Sequence 19, Appl	202	4	40.0	33	5	PCT-US95-10398-241	Sequence 241, App
130	4	40.0	16	2	US-08-102-385B-35	Sequence 35, Appl	203	4	40.0	35	4	US-09-001-984C-11	Sequence 11, Appl
131	4	40.0	16	3	US-08-705-875A-14	Sequence 14, Appl	204	4	40.0	35	4	US-09-001-984C-42	Sequence 42, Appl
132	4	40.0	16	3	US-08-602-999A-179	Sequence 179, App	205	4	40.0	37	1	US-08-463-660-8	Sequence 8, Appl
133	4	40.0	19	1	US-07-805-437-5	Sequence 5, Appl	206	4	40.0	37	1	US-08-678-280-8	Sequence 8, Appl
134	4	40.0	19	1	US-08-170-596-12	Sequence 12, Appl	207	4	40.0	38	2	US-08-146-028-18	Sequence 18, Appl
135	4	40.0	19	5	PCT-US91-02942-67	Sequence 67, Appl	208	4	40.0	38	4	US-08-723-425A-18	Sequence 18, Appl
136	4	40.0	20	1	US-08-218-025A-72	Sequence 72, Appl	209	4	40.0	38	4	US-09-112-006-18	Sequence 18, Appl
137	4	40.0	20	1	US-08-170-596-8	Sequence 8, Appl	210	4	40.0	39	2	US-08-455-625-3	Sequence 3, Appl
138	4	40.0	20	1	US-08-373-776-1	Sequence 1, Appl	211	4	40.0	39	2	US-08-455-625-29	Sequence 29, Appl
139	4	40.0	21	1	US-08-436-772-59	Sequence 59, Appl	212	4	40.0	39	4	US-08-455-685-3	Sequence 3, Appl
140	4	40.0	21	2	US-08-436-883B-59	Sequence 59, Appl	213	4	40.0	39	4	US-08-455-685-29	Sequence 29, Appl
141	4	40.0	21	2	US-08-323-436A-12	Sequence 12, Appl	214	4	40.0	39	5	PCT-US94-05142-3	Sequence 3, Appl
142	4	40.0	21	2	US-08-847-848-2	Sequence 12, Appl	215	4	40.0	39	5	PCT-US94-05142-29	Sequence 29, Appl
143	4	40.0	21	2	US-08-413-490-12	Sequence 12, Appl	216	4	40.0	42	3	US-08-924-330A-1	Sequence 1, Appl
144	4	40.0	21	2	US-08-530-792D-19	Sequence 19, Appl	217	4	40.0	42	4	US-09-262-773-202	Sequence 202, App
145	4	40.0	21	3	US-09-146-269-12	Sequence 12, Appl	218	4	40.0	48	3	US-08-665-259-6	Sequence 6, Appl
146	4	40.0	21	3	US-09-142-078-2	Sequence 12, Appl	219	4	40.0	48	3	US-08-762-500-6	Sequence 6, Appl
147	4	40.0	21	4	PCT-US95-13037-12	Sequence 12, Appl	220	4	40.0	49	3	US-08-665-259-5	Sequence 5, Appl
148	4	40.0	21	5	PCT-US91-02942-67	Sequence 67, Appl	221	4	40.0	49	3	US-08-762-500-5	Sequence 5, Appl
149	4	40.0	22	1	US-08-116-733-5	Sequence 5, Appl	222	4	40.0	50	5	PCT-US91-02942-46	Sequence 46, Appl
150	4	40.0	22	2	US-08-969-721-6	Sequence 6, Appl	223	4	40.0	50	5	PCT-US91-02942-47	Sequence 47, Appl
151	4	40.0	22	2	US-08-833-546-9	Sequence 9, Appl	224	4	40.0	50	5	PCT-US91-02942-48	Sequence 48, Appl
152	4	40.0	22	4	US-09-388-664-9	Sequence 9, Appl	225	4	40.0	50	5	PCT-US91-02942-49	Sequence 49, Appl
153	4	40.0	23	1	US-07-826-928A-11	Sequence 11, Appl	226	4	40.0	50	5	PCT-US91-02942-50	Sequence 50, Appl
154	4	40.0	23	1	US-08-218-025A-152	Sequence 132, App	227	4	40.0	51	1	US-08-570-157-16	Sequence 16, Appl
155	4	40.0	23	2	US-08-493-235-40	Sequence 40, Appl	228	4	40.0	51	1	US-08-570-157-17	Sequence 17, Appl
156	4	40.0	23	2	US-08-492-027A-3	Sequence 3, Appl	229	4	40.0	52	1	US-08-294-189-17	Sequence 17, Appl
157	4	40.0	23	3	US-08-456-670B-22	Sequence 22, Appl	230	4	40.0	53	1	US-08-974-549A-184	Sequence 184, App
158	4	40.0	23	3	US-08-485-324-15	Sequence 15, Appl	231	4	40.0	53	4	US-08-306-871-25	Sequence 25, Appl
159	4	40.0	23	3	US-08-447-506-15	Sequence 15, Appl	232	4	40.0	59	1	US-08-569-859-25	Sequence 25, Appl
160	4	40.0	23	3	US-08-235-437-15	Sequence 15, Appl	233	4	40.0	59	1	US-08-468-001A-23	Sequence 23, Appl
161	4	40.0	23	4	US-08-447-515-15	Sequence 15, Appl	234	4	40.0	60	1	US-08-447-925-1	Sequence 1, Appl
162	4	40.0	24	1	US-08-488-252-31	Sequence 31, Appl	235	4	40.0	60	1	US-08-099-354-7	Sequence 7, Appl
163	4	40.0	24	2	US-08-407-252-4	Sequence 4, Appl	236	4	40.0	60	2	US-08-288-059-33	Sequence 33, Appl
164	4	40.0	24	3	US-08-705-875A-15	Sequence 15, Appl	237	4	40.0	65	4	US-09-188-930-297	Sequence 297, App
165	4	40.0	25	1	US-08-451-405A-3	Sequence 3, Appl	238	4	40.0	72	2	US-08-535-450-12	Sequence 12, Appl
166	4	40.0	25	2	US-08-474-696A-4	Sequence 4, Appl	239	4	40.0	72	3	US-08-338-579A-106	Sequence 106, App
167	4	40.0	25	2	US-08-455-625-25	Sequence 25, Appl	240	4	40.0	72	3	US-08-338-579A-107	Sequence 107, App
168	4	40.0	25	4	US-08-455-685-25	Sequence 25, Appl	241	4	40.0	76	4	US-08-646-265A-113	Sequence 113, App
169	4	40.0	25	5	PCT-US94-05142-25	Sequence 25, Appl	242	4	40.0	81	4	US-08-469-412A-16	Sequence 16, Appl
170	4	40.0	26	1	US-08-170-596-11	Sequence 11, Appl	243	4	40.0	81	2	US-09-021-715-15	Sequence 15, Appl
171	4	40.0	26	3	US-08-705-875A-16	Sequence 16, Appl	244	4	40.0	82	1	US-08-225-575B-12	Sequence 12, Appl
172	4	40.0	26	3	US-08-335-733D-1	Sequence 1, Appl	245	4	40.0	87	1	US-08-497-312-27	Sequence 27, Appl
173	4	40.0	28	2	US-08-392-816-4	Sequence 4, Appl	246	4	40.0	87	1	US-08-497-312-16	Sequence 16, Appl

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 15, 2001, 12:33:53 ; Search time 31.32 Seconds  
(without alignments)  
6.574 Million cell updates/sec

Title: US-09-372-036-31

Perfect score: 10

Sequence: 1 PVAPQEVKK 10

## Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2.6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2.6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2.6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2.6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2.6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2.6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100.0	10	2	US-08-456-670B-31	Sequence 31, Appl
2	100.0	12	2	US-08-456-670B-30	Sequence 30, Appl
3	100.0	21	1	US-08-127-499A-27	Sequence 27, Appl
4	100.0	21	1	US-08-482-847-27	Sequence 27, Appl
5	100.0	232	2	US-08-456-670B-39	Sequence 39, Appl
6	100.0	478	2	US-08-456-670B-40	Sequence 40, Appl
7	100.0	484	1	US-08-127-499A-26	Sequence 26, Appl
8	100.0	484	1	US-08-482-847-26	Sequence 26, Appl
9	70.0	7	1	US-08-127-499A-28	Sequence 28, Appl
10	70.0	7	1	US-08-482-847-28	Sequence 28, Appl
11	60.0	9	2	US-08-456-670B-26	Sequence 26, Appl
12	60.0	20	2	US-08-456-670B-17	Sequence 17, Appl
13	50.0	11	2	US-08-456-670B-35	Sequence 35, Appl
14	50.0	107	2	US-08-810-572A-4	Sequence 4, Appl
15	50.0	118	1	US-08-497-025-10	Sequence 10, Appl
16	50.0	192	1	US-08-086-428B-57	Sequence 57, Appl
17	50.0	192	1	US-08-440-103-44	Sequence 44, Appl
18	50.0	192	1	US-08-440-542-44	Sequence 44, Appl
19	50.0	192	1	US-08-231-366-44	Sequence 44, Appl
20	50.0	192	1	US-08-440-210-44	Sequence 44, Appl
21	50.0	192	2	US-08-468-570-57	Sequence 57, Appl
22	50.0	192	2	US-08-290-665A-57	Sequence 57, Appl
23	50.0	192	4	US-08-444-818-167	Sequence 167, App
24	50.0	192	5	PCT-US95-10398-57	Sequence 57, Appl
25	50.0	231	2	US-08-969-721-8	Sequence 8, Appl
26	50.0	242	6	5273901-3	Patent No. 5273901
27	50.0	242	6	5482709-2	Patent No. 5482709

28	50.0	293	2	US-08-810-572A-2	Sequence 2, Appl
29	50.0	339	4	US-08-444-818-152	Sequence 152, App
30	50.0	339	4	US-08-444-818-156	Sequence 156, App
31	50.0	346	1	US-07-951-715A-19	Sequence 19, Appl
32	50.0	346	2	US-08-459-448A-19	Sequence 19, Appl
33	50.0	346	3	US-08-459-504B-19	Sequence 19, Appl
34	50.0	346	3	US-08-459-504B-19	Sequence 19, Appl
35	50.0	346	3	US-08-459-504B-19	Sequence 19, Appl
36	50.0	399	2	US-08-839-581A-2	Sequence 2, Appl
37	50.0	399	4	US-09-023-591A-2	Sequence 2, Appl
38	50.0	428	1	US-08-353-550-1	Sequence 1, Appl
39	50.0	428	2	US-08-551-687-1	Sequence 1, Appl
40	50.0	431	3	US-08-807-342B-5	Sequence 5, Appl
41	50.0	445	2	US-08-679-635A-3	Sequence 3, Appl
42	50.0	467	1	US-08-140-104A-2	Sequence 2, Appl
43	50.0	469	1	US-08-353-550-6	Sequence 6, Appl
44	50.0	469	2	US-08-551-687-6	Sequence 6, Appl
45	50.0	470	4	US-09-071-709-12	Sequence 12, Appl
46	50.0	486	3	US-08-889-841-8	Sequence 8, Appl
47	50.0	491	3	US-08-889-841-10	Sequence 10, Appl
48	50.0	498	1	US-08-470-202-59	Sequence 59, Appl
49	50.0	498	1	US-08-470-202-60	Sequence 60, Appl
50	50.0	498	1	US-08-471-770-59	Sequence 59, Appl
51	50.0	498	1	US-08-471-770-60	Sequence 60, Appl
52	50.0	498	2	US-08-468-059-59	Sequence 59, Appl
53	50.0	498	2	US-08-468-059-60	Sequence 60, Appl
54	50.0	591	3	US-08-991-408-4	Sequence 4, Appl
55	50.0	605	2	US-08-687-956A-1	Sequence 1, Appl
56	50.0	655	1	US-07-736-178C-2	Sequence 2, Appl
57	50.0	670	2	US-08-366-547-2	Sequence 2, Appl
58	50.0	739	3	US-09-035-648-24	Sequence 24, Appl
59	50.0	739	4	US-09-001-951-24	Sequence 24, Appl
60	50.0	762	2	US-08-907-166-10	Sequence 10, Appl
61	50.0	788	1	US-08-572-225-1	Sequence 1, Appl
62	50.0	1013	2	US-08-866-650-5	Sequence 5, Appl
63	50.0	1013	3	US-09-021-287-5	Sequence 5, Appl
64	50.0	1013	2	US-08-991-408-2	Sequence 2, Appl
65	50.0	1021	1	US-08-497-025-3	Sequence 3, Appl
66	50.0	1187	1	US-08-320-559-28	Sequence 28, Appl
67	50.0	1187	3	US-08-545-860D-28	Sequence 28, Appl
68	50.0	1187	5	PCT-US94-04496-28	Sequence 28, Appl
69	50.0	1210	1	US-08-320-559-26	Sequence 26, Appl
70	50.0	1210	3	PCT-US94-04496-26	Sequence 26, Appl
71	50.0	1210	5	US-08-682-517-15	Sequence 15, Appl
72	50.0	1222	2	US-08-158-223-8	Sequence 8, Appl
73	50.0	1245	1	US-08-304-628-8	Sequence 8, Appl
74	50.0	1245	2	US-08-611-928-8	Sequence 8, Appl
75	50.0	1245	3	US-09-173-891-8	Sequence 8, Appl
76	50.0	1252	2	US-08-682-517-9	Sequence 9, Appl
77	50.0	1313	2	US-08-244-537-2	Sequence 2, Appl
78	50.0	1566	2	US-08-687-956A-23	Sequence 23, Appl
79	50.0	1627	1	US-07-665-792E-9	Sequence 9, Appl
80	50.0	1627	1	US-09-091-219-25	Sequence 25, Appl
81	50.0	2232	4	US-08-417-089-6	Sequence 6, Appl
82	50.0	2247	3	US-08-695-651-6	Sequence 6, Appl
83	50.0	2325	4	US-08-930-285-6	Sequence 6, Appl
84	50.0	2325	4	US-08-276-967-2	Sequence 2, Appl
85	50.0	2476	2	US-08-973-462-28	Sequence 28, Appl
86	50.0	7	1	US-08-261-206A-18	Sequence 18, Appl
87	40.0	7	1	US-09-142-078-11	Sequence 11, Appl
88	40.0	8	2	US-08-769-745-11	Sequence 11, Appl
89	40.0	8	3	US-08-335-733D-38	Sequence 38, Appl
90	40.0	8	3	US-08-335-733D-39	Sequence 39, Appl
91	40.0	8	3	US-08-335-733D-40	Sequence 40, Appl
92	40.0	8	3	US-08-335-733D-41	Sequence 41, Appl
93	40.0	8	3	US-08-335-733D-42	Sequence 42, Appl
94	40.0	9	1	US-08-200-900A-22	Sequence 22, Appl
95	40.0	9	1	US-08-338-634-7	Sequence 7, Appl
96	40.0	9	1	US-08-787-547-77	Sequence 77, Appl
97	40.0	9	1	US-08-335-733D-5	Sequence 5, Appl
98	40.0	9	3	US-09-001-984C-17	Sequence 17, Appl
99	40.0	9	4		
100	40.0	9	4		



Db 83 VAPRO 87

RESULT 50

084412  
ID 084412 PRELIMINARY; PRT; 124 AA.  
AC 084412;  
DT 01-NOV-1998 (TREMblrel. 08, Created)  
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)  
DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)  
DE DNK SUPPRESSOR.  
GN DKSA.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=D/UM-3/CX;  
RX MEDLINE=99000809; Pubmed=9784136;  
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
RA Davis R.W.;  
RT "Genome sequence of an obligate intracellular pathogen of humans:  
RT Chlamydia trachomatis.";  
RL Science 282:754-759(1998).  
DR EMBL: AE001314; AAC68004.1; -  
DR InterPro: IPR000962; -  
DR Pfam: PF01258; zf\_dskA\_trar; 1.  
DR PRINTS: PR00618; DKSANFINGER.  
SQ SEQUENCE 124 AA; 13938 MW; 015FBDS53B65733 CRC64;

Query Match 50.0%; Score 5; DB 2; length 124;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 QEVKK 10  
|||  
Db 31 QEVKK 35

Search completed: August 15, 2001, 12:42:51  
Job time: 536 sec

Wed Aug 15 13:35:26 2001

Db 100 TOEVK 104

RESULT 47

09SVK9 PRELIMINARY: PRT: 114 AA.

AC 09SVK9;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE HYPOTHETICAL 12.9 KDA PROTEIN.

GN F1B83.180.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsis.

NCBI\_TaxID=3702;

OX (1)

RP SEQUENCE FROM N.A.

RA Rieger M., Gabel C., Mueller-Auer S., Schaefer M., Zipp M.,

RA Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C., Queciet F.,

RA Salanoubat M.,

RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.

RM (12)

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (Jun-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AL049862; CAB42919.1; -.

SK Hypothetical protein.

QY SEQUENCE 114 AA; 12923 MW; F8156E1465887DBB CRC64;

Query Match 50.0%; Score 5; DB 10; Length 114;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 OEYK 10

|||||

Db 43 OEYK 47

RESULT 48

ID 09N3V1 PRELIMINARY: PRT: 121 AA.

AC 09N3V1;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE HYPOTHETICAL PROTEIN Y47D7A.H.

GN Y47D7A.H.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Felodierinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RM (1)

RP SEQUENCE FROM N.A.

RA STRAIN-BRISTOL N2;

RC MEDLINE-99069613; PubMed-9851916;

RA None;

RT "Genome sequence of the nematode C. elegans: a platform for

RT investigating biology. The C. elegans Sequencing Consortium.";

RL Science 283:2012-2018(1998).

RM (12)

RP SEQUENCE FROM N.A.

RA STRAIN-BRISTOL N2;

RC Waterston R.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC024790; AAF60636.1; -.

SO SEQUENCE 121 AA; 13441 MW; FF21FD197434C60F CRC64;

Query Match 50.0%; Score 5; DB 5; Length 121;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PVAPT 5

|||||

Db 46 PVAPT 50

RESULT 49

09VYZ8 PRELIMINARY: PRT: 122 AA.

AC 09VYZ8;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE CG11750 PROTEIN.

GN CG11750.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI\_TaxID=7227;

OX (1)

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RM MEDLINE-20196006; PubMed-10731132;

RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sulton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Bailev R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,

RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,

RA Jalali M., Kalush F., Kapran G.H., Ke Z., Kennison J.A., Kethum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,

RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Moberg C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Put V., Reese M.G.,

RA Reinher K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL: AE003485; AAF48034.1; -.

DR Flybase: FBgn0030294; CG11750.

SO SEQUENCE 122 AA; 13243 MW; D7AB3CA104BCID9D CRC64;

Query Match 50.0%; Score 5; DB 5; Length 122;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VAPTQ 6

RESULT 43  
 09WK87 PRELIMINARY; PRT; 113 AA.  
 AC 09WK87;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE P24 CORE PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K88, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
 RA Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
 RT "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda."  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ228278; CAA12887.1; -  
 DR InterPro: IPR000721; -  
 DR Pfam: PF00607; ga9\_p24; 1.  
 FT NON\_TER 1  
 FT 113  
 SO SEQUENCE 113 AA; 12856 MW; 60D790BE02A1F9AF CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEYK 9  
 11111  
 DB 100 TOEYK 104

RESULT 44  
 09WK86 PRELIMINARY; PRT; 113 AA.  
 ID 09WK86;  
 AC 09WK86;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE P24 CORE PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K101, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
 RA Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
 RT "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda."  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ228280; CAA12888.1; -  
 DR InterPro: IPR000721; -  
 DR Pfam: PF00607; ga9\_p24; 1.  
 FT NON\_TER 1  
 FT 113  
 SO SEQUENCE 113 AA; 12953 MW; 3883B5A94D5F3227 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEYK 9  
 11111  
 DB 100 TOEYK 104

RESULT 45  
 09WK47 PRELIMINARY; PRT; 113 AA.  
 ID 09WK47;  
 AC 09WK47;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE P24 CORE PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K96, CLADE D;  
 RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
 RA Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
 RT "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda."  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ228320; CAA12927.1; -  
 DR InterPro: IPR000721; -  
 DR Pfam: PF00607; ga9\_p24; 1.  
 FT NON\_TER 1  
 FT 113  
 SO SEQUENCE 113 AA; 13047 MW; 8B74BB2619F608A3 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEYK 9  
 11111  
 DB 100 TOEYK 104

RESULT 46  
 09WK45 PRELIMINARY; PRT; 113 AA.  
 ID 09WK45;  
 AC 09WK45;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE P24 CORE PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K100, CLADE D;  
 RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
 RA Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
 RT "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda."  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ228322; CAA12929.1; -  
 DR InterPro: IPR000721; -  
 DR Pfam: PF00607; ga9\_p24; 1.  
 FT NON\_TER 1  
 FT 113  
 SO SEQUENCE 113 AA; 13060 MW; 6789B057FC791F27 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEYK 9  
 11111

09WK93  
ID 09WK93: PRELIMINARY: PRT: 113 AA.  
AC 09WK93:  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE P24 CORE PROTEIN (FRAGMENT).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K78, CLADE A;  
RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
RT Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
RT "Molecular epidemiology of HIV-1 in a rural community in south west  
RT Uganda."  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ228272; CA12881.1; --  
DR InterPro: IPR000721; --  
DR Pfam: PF00607; gag\_p24; 1.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 113 AA; 12886 MW; 3CD5A0F417EDEFFA CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 5 TOEVK 9  
DB 100 TOEVK 104

RESULT 40  
09WK91  
ID 09WK91: PRELIMINARY: PRT: 113 AA.  
AC 09WK91:  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE P24 CORE PROTEIN (FRAGMENT).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K81, CLADE A;  
RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
RT Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
RT "Molecular epidemiology of HIV-1 in a rural community in south west  
RT Uganda."  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ228274; CA12883.1; --  
DR InterPro: IPR000721; --  
DR Pfam: PF00607; gag\_p24; 1.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 113 AA; 12928 MW; E93F1F9A75D0408A CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 5 TOEVK 9  
DB 100 TOEVK 104

RESULT 41  
09WK90  
ID 09WK90: PRELIMINARY: PRT: 113 AA.  
AC 09WK90:  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE P24 CORE PROTEIN (FRAGMENT).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K82, CLADE A;  
RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
RT Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
RT "Molecular epidemiology of HIV-1 in a rural community in south west  
RT Uganda."  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ228275; CA12884.1; --  
DR InterPro: IPR000721; --  
DR Pfam: PF00607; gag\_p24; 1.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 113 AA; 12928 MW; 993ED8EA7185148B CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 5 TOEVK 9  
DB 100 TOEVK 104

RESULT 42  
09WK88  
ID 09WK88: PRELIMINARY: PRT: 113 AA.  
AC 09WK88:  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE P24 CORE PROTEIN (FRAGMENT).  
GN GAG.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K87, CLADE A;  
RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
RT Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
RT "Molecular epidemiology of HIV-1 in a rural community in south west  
RT Uganda."  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ228277; CA12886.1; --  
DR InterPro: IPR000721; --  
DR Pfam: PF00607; gag\_p24; 1.  
FT NON\_TER 1  
FT NON\_TER 1  
SQ SEQUENCE 113 AA; 12890 MW; A12E6EC8AEC35310 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 5 TOEVK 9  
DB 100 TOEVK 104

AC 09WK1: 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE P24 CORE PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K29, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemwera A., Lyagoba F., Morgan D.,  
 RA Duffield M., Bityahwaho B., Whitworth J., Oram J.;  
 RT "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda."  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AJ228262; CAA12872.1;  
 DR EMBL: AJ228262; CAA12872.1;  
 DR InterPro: IPR000721;  
 DR Pfam: PF00607; gag\_p24; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SO SEQUENCE 113 AA; 12932 MW; 3710BEDF76F0F193 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVR 9  
 |||||  
 DB 100 TOEVR 104

RESULT 36  
 09WK98 PRELIMINARY; PRT: 113 AA.  
 AC 09WK98: 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE P24 CORE PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K54, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemwera A., Lyagoba F., Morgan D.,  
 RA Duffield M., Bityahwaho B., Whitworth J., Oram J.;  
 RT "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda."  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AJ228267; CAA12876.1;  
 DR EMBL: AJ228267; CAA12876.1;  
 DR InterPro: IPR000721;  
 DR Pfam: PF00607; gag\_p24; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SO SEQUENCE 113 AA; 12946 MW; 04C146666476C072 CRC64;

Query Match -50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVR 9  
 |||||  
 DB 100 TOEVR 104

RESULT 37  
 09WK96

ID 09WK96 PRELIMINARY; PRT: 113 AA.  
 AC 09WK96: 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE P24 CORE PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K73, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemwera A., Lyagoba F., Morgan D.,  
 RA Duffield M., Bityahwaho B., Whitworth J., Oram J.;  
 RT "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda."  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AJ228269; CAA12878.1;  
 DR EMBL: AJ228269; CAA12878.1;  
 DR InterPro: IPR000721;  
 DR Pfam: PF00607; gag\_p24; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SO SEQUENCE 113 AA; 12989 MW; 1A192C3958851E90 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVR 9  
 |||||  
 DB 100 TOEVR 104

RESULT 38  
 09WK94 PRELIMINARY; PRT: 113 AA.  
 AC 09WK94: 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE P24 CORE PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K75, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemwera A., Lyagoba F., Morgan D.,  
 RA Duffield M., Bityahwaho B., Whitworth J., Oram J.;  
 RT "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda."  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AJ228271; CAA12880.1;  
 DR EMBL: AJ228271; CAA12880.1;  
 DR InterPro: IPR000721;  
 DR Pfam: PF00607; gag\_p24; 1.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SO SEQUENCE 113 AA; 12956 MW; 8A19AE2BC185118F CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVR 9  
 |||||  
 DB 100 TOEVR 104

RESULT 39  
 09WK96

OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
 RT Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
 "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda";  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO GAG GENE PROTEIN P24 (CORE NUCLEOCAPSID PROTEIN).  
 DR EMBL: AJ228259; CAI2870.1; -  
 DR InterPro: IPR000721; -  
 DR Pfam: PF00607; gag\_p24; 1.  
 KM Core protein; Polyprotein.  
 FT NON\_TER 1 113  
 FT SEQUENCE 113 AA; 12826 MW; FD2B3B6BE6958BE4 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVR 9  
 |||||  
 DB 100 TOEVR 104

RESULT 32  
 O93053  
 ID 093053 PRELIMINARY; PRT; 113 AA.  
 AC O93053;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 GN GAG POLYPROTEIN [CONTRAINS: CORE PROTEIN(S) P24] (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K38, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
 RT Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
 "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda";  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO GAG GENE PROTEIN P24 (CORE NUCLEOCAPSID PROTEIN).  
 DR EMBL: AJ228263; CAI2874.1; -  
 DR InterPro: IPR000721; -  
 DR Pfam: PF00607; gag\_p24; 1.  
 KM Core protein; Polyprotein.  
 FT NON\_TER 1 113  
 FT SEQUENCE 113 AA; 13035 MW; 35EB47CDEB22A620 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVR 9  
 |||||  
 DB 100 TOEVR 104

RESULT 33  
 O9WKA3 PRELIMINARY; PRT; 113 AA.  
 ID O9WKA3;  
 AC O9WKA3;

DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE P24 CORE PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K11, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
 RT Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
 "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda";  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC EMBL: AJ228258; CAI2869.1; -  
 DR InterPro: IPR000721; -  
 DR Pfam: PF00607; gag\_p24; 1.  
 FT NON\_TER 1 113  
 FT SEQUENCE 113 AA; 12919 MW; 87EB5C54D76C063 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVR 9  
 |||||  
 DB 100 TOEVR 104

RESULT 34  
 O9WKA2  
 ID O9WKA2 PRELIMINARY; PRT; 113 AA.  
 AC O9WKA2;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 GN P24 CORE PROTEIN (FRAGMENT).  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K13, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
 RT Duffield M., Biryahwaho B., Whitworth J., Oram J.;  
 "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda";  
 RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC EMBL: AJ228260; CAI2871.1; -  
 DR InterPro: IPR000721; -  
 DR Pfam: PF00607; gag\_p24; 1.  
 FT NON\_TER 1 113  
 FT SEQUENCE 113 AA; 12878 MW; 03972190189E3FD3 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVR 9  
 |||||  
 DB 100 TOEVR 104

RESULT 35  
 O9WKA1 PRELIMINARY; PRT; 113 AA.  
 ID O9WKA1

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HCV-BB8;  
 RA Somsavilai S., Kanlstanon D.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
 DR EMBL: U23395; AAA64861.1; -.  
 DR InterPro: IPR002519; -.  
 DR Pfam: PF01539; HCV\_env; 1.  
 DR Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 KW Transmembrane.  
 FT NON\_TER 1 1  
 SO SEQUENCE 102 AA; 11040 MW; 2C560825E0AD043E CRC64;

Query Match 50.0%; Score 5; DB 14; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
 |||||  
 DB 24 PVAPT 28

RESULT 28  
 ID 091E25 PRELIMINARY; PRT; 108 AA.  
 AC 091E25;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN GP105 (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 2.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11709;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Parreira R., Esteves A., Piedade J., Venenno T., Canas-Ferreira F.;  
 RT "Inter-individual genetic variability of immunodeficiency virus type 2  
 isolates from Guinea-Bissau (West-Africa)";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ246878; CAB95055.1; -.  
 DR InterPro: IPR000777; -.  
 DR Pfam: PF00516; GP120; 1.  
 DR NON\_TER 1 1  
 FT NON\_TER 108 108  
 SO SEQUENCE 108 AA; 12547 MW; FC22486D50C03680 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVK 9  
 |||||  
 DB 72 TOEVK 76

RESULT 29  
 ID 09UG25 PRELIMINARY; PRT; 112 AA.  
 AC 09UG25;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE HYPOHETICAL. 12.8 KDA PROTEIN.  
 GN KRF2P564A122.  
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RA Duesterhoeft A., Lauber J., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL110269; CAB53708.1; -.  
 DR InterPro: IPR001496; -.  
 DR InterPro: IPR001680; -.  
 DR SMART: SM00320; WD40; 1.  
 KW Hypothetical protein.  
 SO SEQUENCE 112 AA; 12790 MW; 76795A2E6D1112B5 CRC64;

Query Match 50.0%; Score 5; DB 4; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PTOEV 8  
 |||||  
 DB 91 PTOEV 95

RESULT 30  
 ID 093051 PRELIMINARY; PRT; 113 AA.  
 AC 093051;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GAG POLYPROTEIN [CONTAINS: CORE PROTEIN(S) P24] (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KI, CLADE A;  
 RA Kaleebu P., Hamilton L., Rutebemberwa A., Lyagoba F., Morgan D.,  
 RA Dufield M., Bityawho B., Whitworth J., Oram J.;  
 RT "Molecular epidemiology of HIV-1 in a rural community in south west  
 Uganda";  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO GAG GENE PROTEIN P24 (CORE NUCLEOCAPSID PROTEIN).  
 DR EMBL: AJ228255; CAAL2866.1; -.  
 DR InterPro: IPR000721; -.  
 DR Pfam: PF00607; gag\_P24; 1.  
 DR Core protein; Polyprotein.  
 KW NON\_TER 1 1  
 FT NON\_TER 113 113  
 SO SEQUENCE 113 AA; 12844 MW; 161CB3BCB10CB CRC64;

Query Match 50.0%; Score 5; DB 14; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVK 9  
 |||||  
 DB 100 TOEVK 104

RESULT 31  
 ID 093052 PRELIMINARY; PRT; 113 AA.  
 AC 093052;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GAG POLYPROTEIN [CONTAINS: CORE PROTEIN(S) P24] (FRAGMENT).  
 GN GAG.

RP elgens.":  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Du Z., Gattung S.:  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Waterston R.:  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U01277; AAA82477.1;  
 SQ SEQUENCE 82 AA; 9058 MW; E6C6AC722EBB125B6 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAPTO 6  
 DB 40 VAPTO 44

RESULT 24  
 Q9DR28 PRELIMINARY; PRT; 88 AA.  
 AC Q9DR28:  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DD88472;  
 RX MEDLINE=21002580; PubMed=11118076;  
 RA Abebe A., Pollakis G., Fontanet A.L., Fisseha B., Tegbaru B.,  
 RA Kiphus A., Testaye G., Negassa H., Cornelissen M., Goudsmit J.,  
 RA Rinde de Wit T.F.:  
 RT "Identification of a genetic sub-cluster of HIV-1 subtype C (C')  
 RT widespread in Ethiopia.";  
 RL AIDS Res. Hum. Retroviruses 16:1909-1914(2000).  
 DR EMBL: AF245564; AAC40677.1;  
 FT NON\_TER 1  
 FT NON\_TER 88  
 SQ SEQUENCE 88 AA; 10146 MW; 501A630530DC4FEB CRC64;

Query Match 50.0%; Score 5; DB 14; Length 88;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 QEVKK 10  
 DB 73 QEVKK 77

RESULT 25  
 Q9WJ74 PRELIMINARY; PRT; 91 AA.  
 AC Q9WJ74:  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GAG PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=TW396;  
 RA Chang K.S.S., Lin C.-I., Salminen M.O., Liao S.-K., Wu A.M.,  
 RA Lin H.-C., Lin R.-Y., Twu S.-C.:  
 RT "HIV type 1 env gene diversity detected in Taiwan.";  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U73091; AAD00206.1;  
 DR InterPro: IPR000721;  
 DR Pfam: PF00607; gag\_p24; 1.  
 FT NON\_TER 1  
 FT NON\_TER 91  
 SQ SEQUENCE 91 AA; 10492 MW; 21A08BBD402AF874 CRC64;

Query Match 50.0%; Score 5; DB 14; Length 91;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVK 9  
 DB 60 TOEVK 64

RESULT 26  
 Q68302 PRELIMINARY; PRT; 95 AA.  
 AC Q68302:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GENOME POLYPROTEIN (CONTAINS: ENVELOPE GLYCOPROTEIN E1 (GP32) (GP35))  
 DE (FRAGMENT).  
 GN EL.  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HCY-BB43;  
 RA Songshivai S., Kanistanon D.:  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
 DR EMBL: U23396; AAB64862.1;  
 DR InterPro: IPR002519;  
 DR Pfam: PF01539; HCY\_env; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 KW Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 95  
 SQ SEQUENCE 95 AA; 10305 MW; 86270AA69397533A CRC64;

Query Match 50.0%; Score 5; DB 14; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
 DB 23 PVAPT 27

RESULT 27  
 Q68301 PRELIMINARY; PRT; 102 AA.  
 AC Q68301:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GENOME POLYPROTEIN (CONTAINS: ENVELOPE GLYCOPROTEIN E1 (GP32) (GP35))  
 DE (FRAGMENT).  
 GN EL.  
 OS Hepatitis C virus.



OY 3 APTOE 7  
|||||  
DB 19 APTOE 23

RESULT 20  
09GL38 PRELIMINARY; PRT; 52 AA.

AC 09GL38: 09GL38: PRELIMINARY; PRT; 52 AA.  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE CALPASTATIN (FRAGMENT).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chung H.Y., Davis M.E., Hines H.C.;  
RT "PCR-SSCP analysis of the bovine calpastatin gene domain L region."  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY008267; AAC33869.1;  
FT NON\_TER 1 1  
FT 52 52  
SQ SEQUENCE 52 AA; 5749 MW; CBD5A7449AFDDA89 CRC64;

Query Match 50.0%; Score 5; DB 6; Length 52;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVK 9  
|||||  
DB 4 TOEVK 8

RESULT 21

09RRH1 PRELIMINARY; PRT; 75 AA.

AC 09RRH1: 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL 7.7 KDA PROTEIN.  
GN DR2520.  
OS Deinococcus radiodurans.  
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RI;  
RA MEDLINE=20036896; PubMed=10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,  
RA Makarova K.S., Aravind L., Minton K.W., Fleischmann R.D.,  
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
radiodurans R1."  
RL Science 286:1571-1577(1999).  
DR EMBL; AE002081; AAF12064.1;  
DR TIGR; DR2520;  
KW Hypothetical protein.  
SQ SEQUENCE 75 AA; 7704 MW; E8E47B7BBAE33B0E CRC64;

Query Match 50.0%; Score 5; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 VAPTO 6  
|||||  
DB 24 VAPTO 28

RESULT 22  
091692 PRELIMINARY; PRT; 77 AA.

AC 091692: 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
GN ENV.  
OS Human immunodeficiency virus type 2.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11709;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lettner T., Norrgren H., Marquina S., Aaby P., Melbye M.,  
RA Poulsen A.-G., Larsen O., Dias F., Escanilla D., Albert J.,  
RA Nauci A. Jr.;  
RT "HIV-2 genetic variation and DNA load in asymptomatic carriers and  
AIDS cases in Guinea-Bissau".  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF023894; AAC63160.1;  
DR InterPro; IPR000777;  
DR Pfam; PF00516; GP120; 1.  
KW Envelope protein.  
FT NON\_TER 1 1  
FT 77 77  
SQ SEQUENCE 77 AA; 9063 MW; 9970B030D95EC9EC CRC64;

Query Match 50.0%; Score 5; DB 14; Length 77;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 OEVRK 10  
|||||  
DB 59 OEVRK 63

RESULT 23

017728 PRELIMINARY; PRT; 82 AA.

AC 017728: 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE C06E4.2 PROTEIN.  
GN C06E4.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabdilitidae;  
OC Rhabdilitidae; Pelodermidae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copesey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
RA Jones M., Kersey J., Kirsten J., Laister N., Latelle P.,  
RA Lightnig J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,  
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
RA Watson A., Weinstock L., Wilkinson-Spratt J., Wohlman P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

RX MEDLINE-96135216; PubMed-8551561;  
 RA Lochelt M., Flugel R.M.;  
 RT "The human foamy virus pol gene is expressed as a Pro-pol polyprotein  
 and not as a Gag-Pol fusion protein.";  
 RL J. Virol. 70:1033-1040(1996).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97126112; PubMed-8971036;  
 RA Boden J., Lochelt M., Winkler I., Flower R.P., Dellus H.,  
 RA Flugel R.M.;  
 RT "Characterization of the spliced pol transcript of feline foamy virus:  
 the splice acceptor site of the pol transcript is located in gag of  
 foamy viruses.";  
 RT J. Virol. 70:9024-9027(1996).  
 RL [10]  
 RN SEQUENCE FROM N.A.  
 RA Flugel R.M.;  
 RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U21247; AAB48113.1; -.  
 KW Envelope protein.  
 SO SEQUENCE 989 AA; 113890 MW; E0E8338ECE44E0A5 CRC64;

Query Match 60.0%; Score 6; DB 14; Length 989;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVKK 10  
 Db 185 TOEVKK 190

RESULT 17  
 ID 025772 PRELIMINARY; PRT; 1230 AA.  
 AC 025772;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE HYPOPHETICAL 131.8 KDA PROTEIN.  
 GN HP1157.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OX NCBI\_TaxID-210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-26695 / ATCC 700392;  
 RX MEDLINE-97394467; PubMed-9252185;  
 RA Tomb J.-F., White O., Kervavage A.R., Clayton R.A., Sutton G.G.,  
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
 RA Loftus B., Richardson D., Dodson R., Khairi H.G., Glick A.,  
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
 RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,  
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,  
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the gastric pathogen Helicobacter  
 pylori.";  
 RL Nature 388:539-547(1997).  
 DR EMBL; AE000621; AAD08205.1; -.  
 DR TIGR; HP1157; -.  
 DR InterPro; IPR002718; -.  
 DR Pfam; PF01856; HP\_OMP; 1.  
 KW Hypothetical protein.  
 SO SEQUENCE 1230 AA; 131847 MW; 4BF1506F1E072A10 CRC64;

Query Match 60.0%; Score 6; DB 2; Length 1230;  
 Best Local Similarity 100.0%; Pred. No. 11e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVKK 10  
 Db 173 TOEVKK 178

RESULT 18  
 ID 091BG3 PRELIMINARY; PRT; 1653 AA.  
 AC 091BG3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE PAA.  
 GN PAA.  
 OS Streptococcus criceti.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID-1335;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-E49;  
 RA Tamura H., Kato H.;  
 RT "Cell surface antigen I/II - Streptococcus cricetus.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB042239; BAA95000.1; -.  
 DR InterPro; IPR001899; -.  
 DR InterPro; IPR002965; -.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR PRINTS; PR01217; PRICHTEXTNS.  
 DR PROSITE; PS00343; GRAM\_POS\_ANCHORING; UNKNOWN\_1.  
 SO SEQUENCE 1653 AA; 180862 MW; 93F28E215BF598DE CRC64;

Query Match 60.0%; Score 6; DB 2; Length 1653;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPTQ 6  
 Db 955 PVAPTQ 960

RESULT 19  
 ID 063334 PRELIMINARY; PRT; 23 AA.  
 AC 063334;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE ALPHA-2-MACROGLOBULIN (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID-10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WISTAR;  
 RX MEDLINE-85207604; PubMed-2581948;  
 RA Northmann W., Heisig M., Kunz D., Heinrich P.C.;  
 RT "Molecular cloning of cdna sequences for rat alpha 2-macroglobulin and  
 RT measurement of its transcription during experimental inflammation.";  
 RL J. Biol. Chem. 260:6200-6205(1985).  
 DR EMBL; M84369; AAAA1594.1; -.  
 DR InterPro; IPR001599; -.  
 DR Pfam; PF00207; A2M; 1.  
 DR NON\_TER 1  
 FT NON\_TER 23  
 SO SEQUENCE 23 AA; 2526 MW; E2D721FBI223876 CRC64;

Query Match 50.0%; Score 5; DB 11; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 60.0%; Score 6; DB 5; Length 819;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPRO 6  
 |||||  
 Db 563 PVAPRO 568

RESULT 14  
 O98830 PRELIMINARY; PRT; 988 AA.  
 ID O98830;  
 AC O98830;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last annotation update)  
 DE PROVIRAL DNA, HSRV1 DELETION VARIANT.  
 GN ENV.  
 OS Human foamy virus.  
 OC Viruses; Retroid viruses; Retroviridae; Spumavirus.  
 OX NCBI\_TaxID=11641;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Schmidt M., Herchenröder O., Heeney J.L., Rehwilms A.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA Rehwilms A.;  
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y07723; CA66894.1; -;  
 DR EMBL; Y07724; CA66900.1; -;  
 DR EMBL; Y07725; CA66904.1; -;  
 SO SEQUENCE 988 AA; 113762 MW; 0E9653BF7D7C6B CRC64;

Query Match 60.0%; Score 6; DB 14; Length 988;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVKK 10  
 |||||  
 Db 185 TOEVKK 190

RESULT 15  
 O87041 PRELIMINARY; PRT; 988 AA.  
 ID O87041;  
 AC O87041;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE SFVCPZ, COMPLETE GENOME.  
 GN ENV.  
 OS Simian foamy virus.  
 OC Viruses; Retroid viruses; Retroviridae; Spumavirus.  
 OX NCBI\_TaxID=11642;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA STRAIN-SFVCPZ;  
 RC MEDLINE=94240804; Pubmed=8184531;  
 RA Herchenröder O., Renne R., Loncar D., Cobb E.K., Murphy K.K.,  
 RA Schneider J., Mergla A., Luciw P.A.;  
 RT "Isolation, cloning, and sequencing of simian foamy viruses from  
 RT chimpanzees (SFVCPZ): high homology to human foamy virus (HFV).";  
 RL Virology 201:187-199(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SFVCPZ;  
 RT Luciw P.A.;  
 RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U04327; AA19979.1; -;

SO SEQUENCE 988 AA; 113352 MW; 9A88951475BEC62 CRC64;

Query Match 60.0%; Score 6; DB 14; Length 988;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVKK 10  
 |||||  
 Db 185 TOEVKK 190

RESULT 16  
 P90288 PRELIMINARY; PRT; 989 AA.  
 ID P90288;  
 AC P90288;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE ENVELOPE PROTEIN.  
 OS Human spumaretrovirus (foamy virus).  
 OC Viruses; Retroid viruses; Retroviridae; Spumavirus.  
 OX NCBI\_TaxID=11963;  
 RN [1]  
 RN SEQUENCE OF 980-989 FROM N.A.  
 RP MEDLINE=88004420; Pubmed=2820721;  
 RA Flugel R.M., Rehwilms A., Maurer B., Darai G.;  
 RT "Nucleotide sequence analysis of the env gene and its flanking regions  
 RT of the human spumaretrovirus reveals two novel genes.";  
 RL EMBO J. 6:2077-2084(1987).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RA Maurer B., Bannert H., Rehwilms A., Darai G., Flugel R.M.;  
 RL (In) Bolognesi D.P. (eds.);  
 RL Human retroviruses, cancer and AIDS, pp.75-88, Alan Liss Inc.,  
 RL New York (1988).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=91101282; Pubmed=1846194;  
 RA "Analysis of splicing patterns of human spumaretrovirus by polymerase  
 RT chain reaction reveals complex RNA structures.";  
 RL J. Virol. 65:727-735(1991).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=91335779; Pubmed=1651600;  
 RA Lochelt M., Zentgraf H., Flugel R.M.;  
 RT "Construction of an infectious DNA clone of the full-length human  
 RT spumaretrovirus genome and mutagenesis of the bel 1 gene.";  
 RL Virology 184:43-54(1991).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=93348264; Pubmed=8394017;  
 RA Lochelt M., Muranyi W., Flugel R.M.;  
 RT "Human foamy virus genome possesses an internal, Bel-1-dependent and  
 RT functional promoter.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7317-7321(1993).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=95009332; Pubmed=7917520;  
 RA Weissenberger J., Flugel R.M.;  
 RT "Identification and characterization of the bel 3 protein of human  
 RT foamy virus.";  
 RL AIDS Res. Hum. Retroviruses 10:595-600(1994).  
 RN [7]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=95380275; Pubmed=7544460;  
 RA Kogel D., Aboud M., Flugel R.M.;  
 RT "Mutational analysis of the reverse transcriptase and ribonuclease H  
 RT domains of the human foamy virus.";  
 RL Nucleic Acids Res. 23:2621-2625(1995).  
 RN [8]  
 RN SEQUENCE FROM N.A.



OC Viruses; Retrovirdae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=96-62333;  
 RA Belida F.J., Barlow K.L., Murphy G., Parry J.V., Clewley J.P.;  
 RT "A dual subtype B/E HIV-1 infection with a novel V3 loop crown motif  
 among infections acquired in South East Asia and imported into  
 RT England";  
 RL Submitted.(FEB-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1 SIMILARITY: TO GAG GENE PROTEIN P24 (CORE NUCLEOCAPSID PROTEIN).  
 DR EMBL; AJ224197; CAA11887.1; -  
 DR Interpro: IPR000721; -  
 DR Pfam: PF00607; gag\_p24; 1.  
 KM Core protein; Polypotein.  
 FT NON\_TER 1 1  
 FT SEQUENCE 174 AA; 19667 MW; D6F18B245B0707DF CRC64;  
 SQ

Query Match 60.0%; Score 6; DB 14; Length 174;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEYK 10  
 |||||  
 Db 154 TOEYK 159

RESULT 8  
 O9KVZ4 PRELIMINARY; PRT; 274 AA.  
 ID O9KVZ4  
 AC O9KVZ4;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE BETA-LACTAMASE PRECURSOR.  
 GN OXA-23.  
 OS Acinetobacter baumannii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
 OC Acinetobacter.  
 OX NCBI\_TaxID=470;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RYC 52763/97;  
 RX MEDLINE=20277880; Pubmed=10817708;  
 RA Bou G., Oliver A., Martinez-Beltran J.;  
 RT "OXA-24, a novel class D beta-lactamase with carbapenemase activity in  
 RT an Acinetobacter baumannii clinical strain";  
 RL Antimicrob. Agents Chemother. 44:1556-1561(2000).  
 DR EMBL; AJ239129; CAB92323.1; -  
 KW Signal.  
 FT SIGNAL 1 7 POTENTIAL.  
 FT SEQUENCE 274 AA; 30810 MW; 5343DC532ABAB4E6 CRC64;  
 SQ

Query Match 60.0%; Score 6; DB 2; Length 274;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEYK 10  
 |||||  
 Db 197 TOEYK 202

RESULT 9  
 ID O26617 PRELIMINARY; PRT; 486 AA.  
 AC O26617;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE SPCF1.

OS Strongylocentrotus purpuratus (Purple sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;  
 OC Strongylocentrotus.  
 OX NCBI\_TaxID=7668;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95301110; Pubmed=7781910;  
 RA Zeller R.W., Coffman J.A., Harrington M.G., Britten R.J.,  
 RA Davidson E.H.;  
 RT "SPCF1, a sea urchin embryo DNA-binding protein, exists as five  
 RT nested variants encoded by a single mRNA";  
 RL Dev. Biol. 169:713-727(1995).  
 DR EMBL; U18784; AAA86479.1; -  
 DR Interpro: IPR000255; -  
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; UNKNOWN.1.  
 SQ SEQUENCE 486 AA; 54616 MW; A9ABF23F8B26357 CRC64;

Query Match 60.0%; Score 6; DB 5; Length 486;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PROEYK 9  
 |||||  
 Db 40 PROEYK 45

RESULT 10  
 O9UNJ0 PRELIMINARY; PRT; 609 AA.  
 ID O9UNJ0  
 AC O9UNJ0;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE POLYCYSTIN-2-LIKE PROTEIN.  
 GN PKD12.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TESTIS;  
 RA Veldhuisen B., Spruit L., Dauwerse H.G., Breuning M.H., Peters D.J.M.;  
 RT "Genes homologous to the autosomal dominant polycystic kidney disease  
 RT 2 (PKD2) gene";  
 RL Submitted.(JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF118125; AAD46478.1; -  
 DR Interpro: IPR000636; -  
 DR Interpro: IPR001682; -  
 DR Interpro: IPR001696; -  
 DR Interpro: IPR001968; -  
 DR Interpro: IPR002111; -  
 DR Pfam: PF00520; ion\_trans; 1.  
 DR PRINTS; PR00170; NACHANNEL.  
 DR PRODOM; PD003549; -; 1.  
 SQ SEQUENCE 609 AA; 71906 MW; D5DCF9829B9A41AB CRC64;

Query Match 60.0%; Score 6; DB 4; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEYK 10  
 |||||  
 Db 298 TOEYK 303

RESULT 11  
 ID O9NZM6 PRELIMINARY; PRT; 624 AA.  
 AC O9NZM6;  
 ID O9NZM6;

Query Match 70.0%; Score 7; DB 14; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAPTOEV 8  
 |||||  
 DB 39 VAPTOEV 45

RESULT 4  
 O9PM45 PRELIMINARY; PRT; 227 AA.  
 AC O9PM45;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE PUTATIVE TONB TRANSPORT PROTEIN.  
 GN TONB2 OR CJI630.  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 OC Campylobacter.  
 OX NCBI\_TaxID=197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCCTC 11168;  
 RX MEDLINE=20150912; PubMed=10688204;  
 RA Parhill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,  
 Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,  
 Jagsall K., Kariyeh A.V., Moule S., Pallen M.J., Penn C.W.,  
 Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,  
 Whitehead S., Barrall B.G.;  
 RA "The genome sequence of the food-borne pathogen Campylobacter jejuni  
 RT reveals hypervariable sequences."  
 RL Nature 403:665-668(2000).  
 DR EMBL: AL139079; CAB73618.1; -.  
 DR InterPro: IPR003538; -.  
 DR PRINTS: PR01374; TONBPROTEIN.  
 SQ SEQUENCE 227 AA; 26298 MW; 7AEFFB148480DCD CRC64;

Query Match 70.0%; Score 7; DB 2; Length 227;  
 Best Local Similarity 100.0%; Pred. No. 2.2;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APPOEVK 9  
 |||||  
 DB 105 APPOEVK 111

RESULT 5  
 O9HYT7 PRELIMINARY; PRT; 102 AA.  
 AC O9HYT7;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE HYPOTHETICAL PROTEIN PA3307.  
 GN PA3307.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 Garner R.L., Colity L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
 Brody L.L., Coulter S.N., Folger K.R., Kas A., Laibig K., Lim R.M.,  
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Ralzer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen."  
 RL Nature 406:959-964(2000).  
 DR EMBL: AE004753; AAC06695.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 102 AA; 11114 MW; D92FE4B8B95286B CRC64;

Query Match 60.0%; Score 6; DB 2; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPTO 6  
 |||||  
 DB 27 PVAPTO 32

RESULT 6  
 O9X9T5 PRELIMINARY; PRT; 116 AA.  
 AC O9X9T5;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
 DE PUTATIVE REGULATORY PROTEIN.  
 GN SC011.16.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Seeger K., Harris D.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA James K.D., Parkhill J., Barrall B.G., Rajandream M.A.;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kleser H.M., Denaplatte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RA "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
 RL MOL. Microbiol. 21:77-96(1996).  
 DR EMBL: AF096823; CAB46971.1; -.  
 SQ SEQUENCE 116 AA; 12228 MW; 4B6986EB74C7B6CD CRC64;

Query Match 60.0%; Score 6; DB 2; Length 116;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVKK 10  
 |||||  
 DB 29 TOEVKK 34

RESULT 7  
 O93041 PRELIMINARY; PRT; 174 AA.  
 AC O93041;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GAG POLYPROTEIN [CONTAINS: CORE PROTEIN(S) P24] (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.

```

969 5 50.0 1566 2 053414 053414 streptococc
970 5 50.0 1566 2 055308 055308 streptococc
971 5 50.0 1569 2 054183 054183 streptococc
972 5 50.0 1576 4 075096 075096 homo sapien
973 5 50.0 1581 4 015065 015065 homo sapien
974 5 50.0 1625 10 008367 008367 zea mays (m
975 5 50.0 1648 4 015058 015058 homo sapien
976 5 50.0 1685 10 043248 043248 zea mays (m
977 5 50.0 1688 5 09V9D0 09V9D0 dtrosophila
978 5 50.0 1698 5 024440 024440 dtrosophila
979 5 50.0 1698 5 024440 024440 dtrosophila
980 5 50.0 1698 5 03V8R9 03V8R9 dtrosophila
981 5 50.0 1723 2 03JMX8 03JMX8 dtrosophila
982 5 50.0 1785 5 025685 025685 plasmodium
983 5 50.0 1798 5 09VUB6 09VUB6 dtrosophila
984 5 50.0 1819 2 092LY0 092LY0 dtrosophila
985 5 50.0 1927 2 025262 025262 helicobacte
986 5 50.0 1963 10 09LXT9 09LXT9 arabidopsis
987 5 50.0 2084 3 09HEC9 09HEC9 neurospora
988 5 50.0 2164 13 09IAR9 09IAR9 gallus gall
989 5 50.0 2247 14 039818 039818 equine rhin
990 5 50.0 2272 5 017329 017329 caenorhabdi
991 5 50.0 2325 10 041743 041743 zea mays (m
992 5 50.0 2404 5 03VE34 03VE34 dtrosophila
993 5 50.0 2541 5 019663 019663 caenorhabdi
994 5 50.0 2652 2 09RA21 09RA21 vibrio mari
995 5 50.0 2768 5 09VC00 09VC00 dtrosophila
996 5 50.0 2785 4 075691 075691 homo sapien
997 5 50.0 2796 2 048926 048926 mycobacteri
998 5 50.0 2936 5 09NKP7 09NKP7 leishmania
999 5 50.0 2977 5 09VAP9 09VAP9 dtrosophila
1000 5 50.0 3016 14 092531 092531 hepatitis c

```

## ALIGNMENTS

```

RESULT 1
09RNU3 PRELIMINARY: PRT: 86 AA.
AC 09RNU3:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE P60 PROTEIN (FRAGMENT).
GN IAP.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP MEDLINE=90256283; PubMed=2111287;
RX Kohler S.; Leimelster-Wachter M.; Chakraborty T.; Lottspeich F.;
RA Goebel W.;
RT "The gene coding for protein p60 of Listeria monocytogenes and its use
RT as a specific probe for Listeria monocytogenes."
RL Infect. Immun. 58:1943-1950(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Park S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179003; AAD55089.1; -.
FT NON-TER 1 1
FT NON-TER 86 86
SQ SEQUENCE 86 AA: 9194 MW: 71F649A817D697F6 CRC64;

```

```

Query Match 100.0%; Score 10; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PVAPTOEVKK 10

```

```

Db 32 PVAPTOEVKK 41

```

```

RESULT 2
003493 PRELIMINARY: PRT: 478 AA.
AC 003493:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PROTEIN P60 PRECURSOR (INVASION-ASSOCIATED PROTEIN).
GN IAP.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93094153; PubMed=1459966;
RA Buber A.; Kuhn M.; Goebel W.; Kohler S.;
RT "Structural and functional properties of the p60 proteins from
RT different Listeria species."
RL J. Bacteriol. 174:8166-8171(1992).
CC -I- FUNCTION: THIS MAJOR EXTRACELLULAR PROTEIN MAY BE INVOLVED IN THE
CC INVASION OF NONPROFESSIONAL PHAGOCYtic CELLS BY LISTERIA.
DR EMBL; M80351; AAA25280.1; -.
DR InterPro: IPR000064; -.
DR InterPro: IPR002482; -.
DR Pfam: PF00877; NUPC_P60; 1.
DR Pfam: PF01476; LysM; 2.
DR SMART; SM00257; LysM; 1.
KW SIGNAL.
FT CHAIN 1 27 BY SIMILARITY.
FT CHAIN 28 478 PROTEIN P60.
SQ SEQUENCE 478 AA: 49997 MW: D46D050507402344 CRC64;

```

```

Query Match 100.0%; Score 10; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PVAPTOEVKK 10

```

```

Db 148 PVAPTOEVKK 157

```

```

RESULT 3
067673 PRELIMINARY: PRT: 94 AA.
AC 067673:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HEXON PROTEIN (FRAGMENT).
OS porcine adenovirus sp.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=47901;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Pring-Akerblom P.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X95630; CAA64884.1; -.
DR HSP; P03277; IDHX.
DR InterPro: IPR000736; -.
DR Pfam; PF01065; Adeno_hexon; 1.
KW Hexon protein.
FT NON-TER 1 1
FT NON-TER 94 94
SQ SEQUENCE 94 AA: 10807 MW: 2646FBB97DDCC9FD CRC64;

```

823	5	50.0	670	6	Q9X762	Q9X162	cercopithec	896	5	50.0	970	4	Q9Y6X9	Q9Y6X9	homo sapien
824	5	50.0	671	5	Q07431	Q07431	bombyx mori	897	5	50.0	970	4	Q9UGX3	Q9UGX3	homo sapien
825	5	50.0	681	4	Q07543	Q07543	homo sapien	898	5	50.0	977	13	Q91925	Q91925	xenopus lae
826	5	50.0	684	4	Q9Y338	Q9Y338	homo sapien	899	5	50.0	982	3	Q74773	Q74773	schizosach
827	5	50.0	687	11	Q83024	Q83024	mus musculus	900	5	50.0	988	5	Q16851	Q16851	pluteia xy
828	5	50.0	688	5	Q9VXB7	Q9VXB7	drosophila	901	5	50.0	995	14	P90248	P90248	human immun
829	5	50.0	689	11	Q88934	Q88934	mus musculus	902	5	50.0	998	1	Q9UYC6	Q9UYC6	pyrococcus
830	5	50.0	698	5	Q01719	Q01719	strongyloce	903	5	50.0	1000	2	Q9WYK5	Q9WYK5	thermotoga
831	5	50.0	691	2	Q9PMJ3	Q9PMJ3	campylobact	904	5	50.0	1002	5	Q9VTC4	Q9VTC4	drosophila
832	5	50.0	691	13	Q57658	Q57658	gallus gall	905	5	50.0	1004	4	Q9UHN7	Q9UHN7	homo sapien
833	5	50.0	700	5	Q96839	Q96839	drosophila	906	5	50.0	1008	13	Q9DER7	Q9DER7	gallus gall
834	5	50.0	700	5	Q9V887	Q9V887	drosophila	907	5	50.0	1013	4	Q43897	Q43897	homo sapien
835	5	50.0	700	5	Q9UAS6	Q9UAS6	drosophila	908	5	50.0	1013	4	Q9NQS4	Q9NQS4	homo sapien
836	5	50.0	703	5	Q9V850	Q9V850	drosophila	909	5	50.0	1021	4	Q93033	Q93033	homo sapien
837	5	50.0	704	5	Q9VA70	Q9VA70	drosophila	910	5	50.0	1022	5	Q61907	Q61907	caenorhabd
838	5	50.0	714	2	Q85056	Q85056	moraxella c	911	5	50.0	1022	13	Q57460	Q57460	brachydanio
839	5	50.0	718	5	Q96721	Q96721	encephalito	912	5	50.0	1025	10	Q9LME3	Q9LME3	arabidopsis
840	5	50.0	728	5	Q9NNK3	Q9NNK3	leishmania	913	5	50.0	1030	4	Q14974	Q14974	homo sapien
841	5	50.0	732	5	Q9Y0C4	Q9Y0C4	leishmania	914	5	50.0	1052	5	Q9W554	Q9W554	drosophila
842	5	50.0	735	13	Q57381	Q57381	xenopus lae	915	5	50.0	1057	5	Q9W294	Q9W294	drosophila
843	5	50.0	739	5	P91060	P91060	caenorhabd	916	5	50.0	1066	4	Q9UGX4	Q9UGX4	homo sapien
844	5	50.0	739	5	Q9N8U6	Q9N8U6	trypanosoma	917	5	50.0	1070	5	Q22796	Q22796	arabidopsi
845	5	50.0	746	5	Q18350	Q18350	caenorhabd	918	5	50.0	1071	10	Q49727	Q49727	arabidopsi
846	5	50.0	748	3	Q9P3D8	Q9P3D8	neurospora	919	5	50.0	1073	10	Q9SRV6	Q9SRV6	arabidopsi
847	5	50.0	750	2	P96256	P96256	mycobacteri	920	5	50.0	1092	2	Q50236	Q50236	zymomonas m
848	5	50.0	763	2	Q9ZBL8	Q9ZBL8	mycobacteri	921	5	50.0	1096	10	Q9LRE5	Q9LRE5	arabidopsi
849	5	50.0	775	2	Q912D2	Q912D2	pseudomonas	922	5	50.0	1113	2	Q9R6S8	Q9R6S8	bacillus ba
850	5	50.0	782	4	Q9U194	Q9U194	homo sapien	923	5	50.0	1115	4	Q9ULLO	Q9ULLO	homo sapien
851	5	50.0	786	6	Q9X5X1	Q9X5X1	bos laurus	924	5	50.0	1127	1	Q58221	Q58221	pyrococcus
852	5	50.0	791	5	Q96454	Q96454	drosophila	925	5	50.0	1132	2	Q9PAB1	Q9PAB1	xylella fas
853	5	50.0	791	5	Q9VW65	Q9VW65	drosophila	926	5	50.0	1181	5	Q9U3A2	Q9U3A2	caenorhabd
854	5	50.0	792	5	Q18174	Q18174	caenorhabd	927	5	50.0	1186	5	Q9U3A3	Q9U3A3	caenorhabd
855	5	50.0	792	5	Q9UNG1	Q9UNG1	ephratista f	928	5	50.0	1199	6	Q28139	Q28139	bos taurus
856	5	50.0	795	5	Q62006	Q62006	branchiosto	929	5	50.0	1212	2	Q9L1C8	Q9L1C8	streptomyce
857	5	50.0	797	4	Q95267	Q95267	homo sapien	930	5	50.0	1217	5	P91457	P91457	caenorhabd
858	5	50.0	797	4	Q9UNN9	Q9UNN9	homo sapien	931	5	50.0	1228	3	Q12754	Q12754	saccharomyc
859	5	50.0	803	4	Q9N2L5	Q9N2L5	homo sapien	932	5	50.0	1251	13	Q91365	Q91365	coturnix co
860	5	50.0	803	4	Q9HBY3	Q9HBY3	homo sapien	933	5	50.0	1262	5	Q9N911	Q9N911	leishmania
861	5	50.0	815	4	Q9P200	Q9P200	homo sapien	934	5	50.0	1268	11	Q63623	Q63623	rattus norv
862	5	50.0	815	10	Q9SHN8	Q9SHN8	arabidopsi	935	5	50.0	1274	10	Q9M680	Q9M680	dunaliella
863	5	50.0	816	4	Q9UFT2	Q9UFT2	homo sapien	936	5	50.0	1276	10	Q81059	Q81059	arabidopsi
864	5	50.0	816	5	Q9VKJ9	Q9VKJ9	drosophila	937	5	50.0	1338	5	Q9CQ08	Q9CQ08	glatridia lam
865	5	50.0	817	3	Q07229	Q07229	saccharomyc	938	5	50.0	1341	4	Q9UFP8	Q9UFP8	homo sapien
866	5	50.0	822	1	Q27154	Q27154	methanobact	939	5	50.0	1357	5	Q9W4M4	Q9W4M4	drosophila
867	5	50.0	831	4	Q9UMS6	Q9UMS6	homo sapien	940	5	50.0	1401	5	Q9VAV5	Q9VAV5	drosophila
868	5	50.0	833	11	Q9EPZ2	Q9EPZ2	mus musculus	941	5	50.0	1415	5	Q61442	Q61442	caenorhabd
869	5	50.0	833	2	Q9S594	Q9S594	mycoplasma	942	5	50.0	1420	5	Q9Y0Y8	Q9Y0Y8	drosophila
870	5	50.0	835	5	Q9V714	Q9V714	drosophila	943	5	50.0	1423	5	Q9W1A0	Q9W1A0	drosophila
871	5	50.0	846	5	Q9V210	Q9V210	drosophila	944	5	50.0	1425	14	P90072	P90072	human immun
872	5	50.0	847	14	Q41537	Q41537	human immun	945	5	50.0	1427	14	Q9WF74	Q9WF74	human immun
873	5	50.0	848	14	Q9QAL9	Q9QAL9	white spot	946	5	50.0	1427	14	Q9WE71	Q9WE71	human immun
874	5	50.0	849	4	Q9NVE4	Q9NVE4	homo sapien	947	5	50.0	1427	14	Q9WF68	Q9WF68	human immun
875	5	50.0	859	14	Q9YU24	Q9YU24	human immun	948	5	50.0	1427	14	Q9WF65	Q9WF65	human immun
876	5	50.0	866	4	Q9H466	Q9H466	homo sapien	949	5	50.0	1429	14	Q9QBE0	Q9QBE0	human immun
877	5	50.0	870	5	Q9GRT4	Q9GRT4	leishmania	950	5	50.0	1430	14	Q89290	Q89290	human immun
878	5	50.0	871	2	Q9PIZ1	Q9PIZ1	campylobact	951	5	50.0	1433	14	P90071	P90071	human immun
879	5	50.0	882	10	Q9LME2	Q9LME2	arabidopsi	952	5	50.0	1433	14	Q9QML4	Q9QML4	human immun
880	5	50.0	887	2	Q917T2	Q917T2	drosophila	953	5	50.0	1433	14	Q91E69	Q91E69	human immun
881	5	50.0	887	2	Q9U212	Q9U212	neisseria m	954	5	50.0	1433	14	Q41791	Q41791	human immun
882	5	50.0	887	2	Q9JU08	Q9JU08	neisseria m	955	5	50.0	1435	14	Q9WLJ3	Q9WLJ3	human immun
883	5	50.0	889	5	Q9V8R8	Q9V8R8	drosophila	956	5	50.0	1440	5	Q27395	Q27395	caenorhabd
884	5	50.0	899	3	Q06132	Q06132	saccharomyc	957	5	50.0	1440	14	Q91C79	Q91C79	chimpanzee
885	5	50.0	904	5	Q9VEX7	Q9VEX7	drosophila	958	5	50.0	1449	14	Q10922	Q10922	caenorhabd
886	5	50.0	908	5	Q9VD00	Q9VD00	drosophila	959	5	50.0	1449	14	Q91080	Q91080	human immun
887	5	50.0	912	2	Q9ZLV5	Q9ZLV5	helicobacte	960	5	50.0	1452	14	Q92888	Q92888	human immun
888	5	50.0	912	5	Q9NHC1	Q9NHC1	drosophila	961	5	50.0	1457	5	Q9VT35	Q9VT35	drosophila
889	5	50.0	912	5	Q9NHB9	Q9NHB9	drosophila	962	5	50.0	1490	5	Q9U485	Q9U485	drosophila
890	5	50.0	916	5	Q9NHC2	Q9NHC2	drosophila	963	5	50.0	1534	6	Q28298	Q28298	canis famli
891	5	50.0	916	5	Q9NHC0	Q9NHC0	drosophila	964	5	50.0					
892	5	50.0	919	5	Q9Z7K3	Q9Z7K3	chlamydia p	965	5	50.0					
893	5	50.0	928	14	Q91MX9	Q91MX9	cercopitlthi	966	5	50.0					
894	5	50.0	969	2	Q9RZ15	Q9RZ15	delnocooc	967	5	50.0					
895	5	50.0	969	4	Q9Y214	Q9Y214	homo sapien	968	5	50.0					



677	5	50.0	492	14	089962	089962 human immun	750	5	50.0	528	10	09SRE2	09sre2 arabisdopsi
678	5	50.0	492	14	09WE75	09WE75 human immun	751	5	50.0	530	2	09I008	09i008 pseudomonas
679	5	50.0	492	14	09WE72	09WE72 human immun	752	5	50.0	535	14	09QMK6	09qmk6 chltia viru
680	5	50.0	492	14	09WF69	09WF69 human immun	753	5	50.0	535	14	09IV39	09iv39 human calic
681	5	50.0	492	14	09WF66	09WF66 human immun	754	5	50.0	537	2	09Z8U8	09z8u8 chlamydia p
682	5	50.0	492	14	09QWR9	09QWR9 human immun	755	5	50.0	537	2	09JRV7	09jrv7 chlamydia p
683	5	50.0	493	10	09EAL3	09EAL3 human immun	756	5	50.0	539	10	080914	080914 arabisdopsi
684	5	50.0	493	10	024326	024326 phaseolus v	757	5	50.0	539	11	060664	060664 mus musculu
685	5	50.0	493	10	09FNR2	09FNR2 arabisdopsi	758	5	50.0	540	4	060230	060230 mus saplien
686	5	50.0	493	10	09FHI9	09FHI9 arabisdopsi	759	5	50.0	542	5	023446	023446 caenorhabdl
687	5	50.0	493	14	09QMR9	09QMR9 human immun	760	5	50.0	543	10	022188	022188 arabisdopsi
688	5	50.0	493	14	09IV95	09IV95 human immun	761	5	50.0	544	11	060764	060764 mus musculu
689	5	50.0	494	14	09I256	09I256 human immun	762	5	50.0	547	13	091349	091349 xenopus lae
690	5	50.0	494	14	09I257	09I257 human immun	763	5	50.0	548	2	09ZEY1	09zey1 listeria mo
691	5	50.0	494	14	09O259	09O259 human immun	764	5	50.0	548	2	09EXF6	09exf6 listeria mo
692	5	50.0	494	14	09OSU0	09OSU0 human immun	765	5	50.0	549	5	024441	024441 drosophila
693	5	50.0	494	14	09OC00	09OC00 human immun	766	5	50.0	552	5	09VI52	09vi52 drosophila
694	5	50.0	494	14	09OBY4	09OBY4 human immun	767	5	50.0	553	2	09RSN4	09rsn4 delnoccocus
695	5	50.0	494	14	09IWA4	09IWA4 human immun	768	5	50.0	557	2	09K869	09k869 aquifex aeo
696	5	50.0	495	14	09OP03	09OP03 human immun	769	5	50.0	557	2	09K926	09k926 bacillus ha
697	5	50.0	495	14	09OB22	09OB22 human immun	770	5	50.0	558	2	09H919	09h919 synecocyst
698	5	50.0	495	14	09IWS3	09IWS3 human immun	771	5	50.0	562	1	09HRL9	09hrl9 halobacteri
699	5	50.0	495	14	076126	076126 human immun	772	5	50.0	567	2	033932	033932 listeria mo
700	5	50.0	496	5	017537	017537 caenorhabdl	773	5	50.0	572	5	026042	026042 penaeus jap
701	5	50.0	496	14	09WSG0	09WSG0 human immun	774	5	50.0	573	3	09B179	09b179 saccharomyc
702	5	50.0	496	14	09WS51	09WS51 human immun	775	5	50.0	573	5	019963	019963 caenorhabdl
703	5	50.0	496	14	09WLS9	09WLS9 human immun	776	5	50.0	574	2	074297	074297 synecocyst
704	5	50.0	496	14	09OST1	09OST1 human immun	777	5	50.0	578	5	09TYL0	09tyl0 caenorhabdl
705	5	50.0	496	14	09DKED	09DKED human immun	778	5	50.0	580	10	09ZVU7	09zvuv7 arabisdopsi
706	5	50.0	497	14	041792	041792 human immun	779	5	50.0	583	5	019777	019777 caenorhabdl
707	5	50.0	497	14	089926	089926 human immun	780	5	50.0	584	4	09UNR7	09unr7 homo saplien
708	5	50.0	497	14	09WLS2	09WLS2 human immun	781	5	50.0	596	10	09LSW2	09lsw2 arabisdopsi
709	5	50.0	497	14	09WLS1	09WLS1 human immun	782	5	50.0	604	2	000509	000509 streptococ
710	5	50.0	497	14	09ID96	09ID96 human immun	783	5	50.0	606	2	053561	053561 streptococ
711	5	50.0	498	14	088150	088150 human immun	784	5	50.0	609	13	013027	013027 xenopus lae
712	5	50.0	498	14	077372	077372 human immun	785	5	50.0	609	14	082670	082670 influenza c
713	5	50.0	498	14	079665	079665 human immun	786	5	50.0	609	14	082670	082670 influenza c
714	5	50.0	498	14	09WIR6	09WIR6 human immun	787	5	50.0	609	14	082671	082671 influenza c
715	5	50.0	498	14	09O239	09O239 human immun	788	5	50.0	609	14	082672	082672 influenza c
716	5	50.0	498	14	09OMM1	09OMM1 human immun	789	5	50.0	609	14	082673	082673 influenza c
717	5	50.0	498	14	09OML3	09OML3 human immun	790	5	50.0	609	14	082674	082674 influenza c
718	5	50.0	498	14	09O650	09O650 human immun	791	5	50.0	609	14	082675	082675 influenza c
719	5	50.0	498	14	09O6H6	09O6H6 human immun	792	5	50.0	620	5	09Y0D4	09y0d4 penaeus mon
720	5	50.0	499	14	09O6G9	09O6G9 human immun	793	5	50.0	624	4	09P2U4	09p2u4 homo saplien
721	5	50.0	498	14	09DKH3	09DKH3 human immun	794	5	50.0	626	4	09P2U3	09p2u3 homo saplien
722	5	50.0	498	14	09DH06	09DH06 human immun	795	5	50.0	627	4	09H172	09h172 homo saplien
723	5	50.0	499	14	072860	072860 human immun	796	5	50.0	633	2	059526	059526 mycoplasma
724	5	50.0	499	14	074834	074834 human immun	797	5	50.0	637	3	008484	008484 saccharomyc
725	5	50.0	499	14	089692	089692 human immun	798	5	50.0	641	5	001667	001667 drosophila
726	5	50.0	499	14	09Q712	09Q712 human immun	799	5	50.0	641	5	09VU01	09vu01 drosophila
727	5	50.0	499	14	09IW66	09IW66 human immun	800	5	50.0	641	14	09IO72	09io72 influenza c
728	5	50.0	500	14	089971	089971 human immun	801	5	50.0	645	5	09NFX7	09nfx7 penaeus van
729	5	50.0	500	14	074806	074806 human immun	802	5	50.0	645	14	09IO73	09io73 influenza c
730	5	50.0	500	14	09YX54	09YX54 human immun	803	5	50.0	645	14	09IO69	09io69 influenza c
731	5	50.0	500	14	09WQ10	09WQ10 human immun	804	5	50.0	646	2	007625	007625 bacillus su
732	5	50.0	502	14	09O6X6	09O6X6 human immun	805	5	50.0	646	14	067417	067417 influenza c
733	5	50.0	502	14	09IDAS	09IDAS human immun	806	5	50.0	647	14	067386	067386 influenza c
734	5	50.0	502	14	09IC78	09IC78 chimpanzee	807	5	50.0	648	14	09IO66	09io66 influenza c
735	5	50.0	503	10	09LD28	09LD28 chimpanzee	808	5	50.0	649	14	039661	039661 influenza v
736	5	50.0	503	14	09LD28	09LD28 chimpanzee	809	5	50.0	649	14	067416	067416 influenza c
737	5	50.0	505	2	025089	025089 human immun	810	5	50.0	649	14	067418	067418 influenza c
738	5	50.0	505	4	09UNR8	09UNR8 homo saplien	811	5	50.0	649	14	067419	067419 influenza c
739	5	50.0	507	4	013064	013064 chimpanzee	812	5	50.0	649	14	09IO71	09io71 influenza c
740	5	50.0	508	6	09QMC8	09QMC8 felis silve	813	5	50.0	649	14	09IO67	09io67 influenza c
741	5	50.0	509	6	09QMC7	09QMC7 bison bison	814	5	50.0	650	14	09IO67	09io67 influenza c
742	5	50.0	512	14	09IO79	09IO79 human immun	815	5	50.0	651	5	019659	019659 caenorhabdl
743	5	50.0	516	10	09SVB0	09SVB0 arabisdopsi	816	5	50.0	662	5	062139	062139 caenorhabdl
744	5	50.0	517	14	071261	071261 human immun	817	5	50.0	667	10	09LJES	09ljies arabisdopsi
745	5	50.0	517	14	073329	073329 human immun	818	5	50.0	668	5	09VWM7	09vwm7 drosophila
746	5	50.0	519	5	09NEU8	09NEU8 caenorhabdl	819	5	50.0	669	5	09VNIO	09vni0 drosophila
747	5	50.0	520	11	035731	035731 mus musculu	820	5	50.0	670	4	075714	075714 homo saplien
748	5	50.0	526	4	09P280	09P280 homo saplien	821	5	50.0	670	4	09UNK7	09unk7 homo saplien
749	5	50.0					822	5	50.0				

531	5	50.0	353	14	Q9ENC6	Q9enc6 human immun	604	5	50.0	452	5	Q27696	Q27696 musca domes
532	5	50.0	353	14	Q9ENC4	Q9enc4 human immun	605	5	50.0	456	11	Q9JTB8	Q9jtb8 mus musculu
533	5	50.0	355	14	Q67685	Q67685 aquifex aeo	606	5	50.0	457	2	Q9ZHC0	Q9zhc0 streptococc
534	5	50.0	359	8	Q37749	Q37749 cephalospor	607	5	50.0	459	14	Q9JRS3	Q9jrs3 human papil
535	5	50.0	360	4	Q9UCM6	Q9ucm6 homo sapien	608	5	50.0	460	5	Q9XTQ9	Q9xtq9 caenorhabd1
536	5	50.0	362	2	Q9KRG9	Q9krg9 yersinia en	609	5	50.0	461	5	Q45051	Q45051 caenorhabd1
537	5	50.0	362	5	Q9NEZ1	Q9nez1 caenorhabd1	610	5	50.0	462	5	Q9W143	Q9w143 drosophila
538	5	50.0	362	10	Q82054	Q82054 saccharum o	611	5	50.0	464	2	P74828	P74828 sphingomona
539	5	50.0	363	10	Q9M6B4	Q9m6b4 vitis vinif	612	5	50.0	465	14	Q90729	Q90729 human papil
540	5	50.0	366	14	Q86313	Q86313 reovirus sp	613	5	50.0	467	4	Q9NXT1	Q9nxt1 homo sapien
541	5	50.0	366	14	Q86314	Q86314 reovirus sp	614	5	50.0	467	14	Q40461	Q40461 human immun
542	5	50.0	366	14	Q86315	Q86315 reovirus sp	615	5	50.0	468	14	Q9ENB2	Q9enb2 human immun
543	5	50.0	366	14	Q86316	Q86316 reovirus sp	616	5	50.0	469	2	Q9ZL71	Q9ztl71 helicobacte
544	5	50.0	366	14	Q86317	Q86317 reovirus sp	617	5	50.0	469	14	Q9IXN9	Q9ixn9 human immun
545	5	50.0	366	14	Q86318	Q86318 reovirus sp	618	5	50.0	469	14	Q9XND9	Q9xnd9 human immun
546	5	50.0	366	14	Q86319	Q86319 reovirus sp	619	5	50.0	469	14	Q9ENC2	Q9enc2 human immun
547	5	50.0	366	14	Q86320	Q86320 reovirus sp	620	5	50.0	469	14	Q9ENC0	Q9enc0 human immun
548	5	50.0	366	14	Q86323	Q86323 reovirus sp	621	5	50.0	469	14	Q9ENB8	Q9enb8 human immun
549	5	50.0	366	14	Q86307	Q86307 reovirus sp	622	5	50.0	470	5	Q19087	Q19087 caenorhabd1
550	5	50.0	366	14	Q86308	Q86308 reovirus sp	623	5	50.0	470	14	Q9EN44	Q9ena4 human immun
551	5	50.0	366	14	Q86309	Q86309 reovirus sp	624	5	50.0	472	14	Q9ENB4	Q9enb4 human immun
552	5	50.0	366	14	Q86310	Q86310 reovirus sp	625	5	50.0	472	14	Q9ENB9	Q9enb9 human immun
553	5	50.0	366	14	Q86311	Q86311 reovirus sp	626	5	50.0	473	2	Q50210	Q50210 synechococc
554	5	50.0	366	14	Q86312	Q86312 reovirus sp	627	5	50.0	473	14	Q9ENB0	Q9enb0 human immun
555	5	50.0	366	14	Q9ENM1	Q9enm1 reovirus (t	628	5	50.0	473	14	Q9ENB8	Q9ena8 human immun
556	5	50.0	368	12	Q87097	Q87097 chimpanzee	629	5	50.0	473	14	Q9EN46	Q9ena6 human immun
557	5	50.0	368	5	Q9F010	Q9f010 pseudomonas	630	5	50.0	473	14	Q9EN97	Q9en97 human immun
558	5	50.0	369	5	Q9N405	Q9n405 caenorhabd1	631	5	50.0	473	14	Q9EN94	Q9en94 human immun
559	5	50.0	378	2	Q55307	Q55307 streptococc	632	5	50.0	474	14	Q9EN90	Q9en90 human immun
560	5	50.0	382	4	Q9H878	Q9h878 homo sapien	633	5	50.0	476	5	Q03996	Q03996 plasmodium
561	5	50.0	385	6	Q9GL14	Q9gl14 microptamo	634	5	50.0	476	14	Q9EN92	Q9en92 human immun
562	5	50.0	385	10	Q9ZP00	Q9zp00 brassica ju	635	5	50.0	477	2	Q9RXT6	Q9rxt6 delnoccocus
563	5	50.0	385	14	Q80627	Q80627 human immun	636	5	50.0	478	5	Q25684	Q25684 plasmodium
564	5	50.0	386	5	P91632	P91632 drosophila	637	5	50.0	481	2	Q9L4Q8	Q9l4q8 clostridium
565	5	50.0	387	4	Q9H8P7	Q9h8p7 homo sapien	638	5	50.0	481	14	Q77697	Q77697 human immun
566	5	50.0	389	2	Q9K7L3	Q9k7l3 bacillus ha	639	5	50.0	481	14	Q9JVI9	Q9jvi9 human immun
567	5	50.0	392	4	Q9NPL5	Q9npl5 homo sapien	640	5	50.0	482	5	Q9XY38	Q9xy38 acanthamoeb
568	5	50.0	393	14	Q41549	Q41549 human immun	641	5	50.0	482	5	Q03998	Q03998 plasmodium
569	5	50.0	394	5	Q9VMU4	Q9vmu4 drosophila	642	5	50.0	484	5	Q9U9P0	Q9u9p0 hydra atten
570	5	50.0	395	2	Q9KE01	Q9ke01 bacillus ha	643	5	50.0	484	14	Q77699	Q77699 human immun
571	5	50.0	398	4	Q9UGL3	Q9ugl3 homo sapien	644	5	50.0	484	14	Q77700	Q77700 human immun
572	5	50.0	400	11	Q9E0U2	Q9equ2 apodemus ag	645	5	50.0	484	14	Q80622	Q80622 human immun
573	5	50.0	407	5	Q9VBM4	Q9vbm4 drosophila	646	5	50.0	486	4	Q9UCX4	Q9ucx4 homo sapien
574	5	50.0	408	4	Q9FD10	Q9fd10 salmonella	647	5	50.0	486	14	Q77804	Q77804 human immun
575	5	50.0	408	4	Q9UGL4	Q9ugl4 homo sapien	648	5	50.0	486	14	Q78566	Q78566 human immun
576	5	50.0	409	2	Q9LKA3	Q9lka3 shigella dy	649	5	50.0	486	14	Q79451	Q79451 human immun
577	5	50.0	410	2	Q05446	Q05446 mycobacteri	650	5	50.0	486	14	Q80628	Q80628 human immun
578	5	50.0	411	14	Q81814	Q81814 hepatitis c	651	5	50.0	486	14	Q80629	Q80629 human immun
579	5	50.0	414	14	Q81329	Q81329 hepatitis c	652	5	50.0	486	14	Q80634	Q80634 human immun
580	5	50.0	414	14	P89958	P89958 hepatitis c	653	5	50.0	487	14	Q80621	Q80621 human immun
581	5	50.0	414	14	P89959	P89959 hepatitis c	654	5	50.0	488	14	Q79362	Q79362 human immun
582	5	50.0	415	14	Q81554	Q81554 hepatitis c	655	5	50.0	489	14	Q79350	Q79350 human immun
583	5	50.0	416	2	Q31444	Q31444 bacillus su	656	5	50.0	489	14	Q79354	Q79354 human immun
584	5	50.0	418	2	Q9K9F3	Q9k9f3 bacillus ha	657	5	50.0	489	14	Q79360	Q79360 human immun
585	5	50.0	418	4	Q75066	Q75066 homo sapien	658	5	50.0	489	14	Q80306	Q80306 human immun
586	5	50.0	418	4	Q9NUA8	Q9nuu8 homo sapien	659	5	50.0	489	14	Q80307	Q80307 human immun
587	5	50.0	421	4	Q9UNN6	Q9unn6 homo sapien	660	5	50.0	489	14	Q80631	Q80631 human immun
588	5	50.0	421	4	Q9UBH9	Q9ubh9 homo sapien	661	5	50.0	489	14	Q80305	Q80305 human immun
589	5	50.0	421	4	Q9Y617	Q9y617 homo sapien	662	5	50.0	490	2	Q9RNV1	Q9rnmv1 bacillus an
590	5	50.0	422	2	P96208	P96208 mycobacteri	663	5	50.0	490	14	Q80623	Q80623 human immun
591	5	50.0	423	2	Q49939	Q49939 mycobacteri	664	5	50.0	490	14	P90073	P90073 human immun
592	5	50.0	434	14	Q89933	Q89933 human immun	665	5	50.0	490	14	P90255	P90255 human immun
593	5	50.0	435	5	Q93322	Q93322 caenorhabd1	666	5	50.0	490	14	Q9WC83	Q9wc83 human immun
594	5	50.0	436	14	Q91V79	Q91v79 human immun	667	5	50.0	491	2	Q9FE18	Q9fe18 alteromonas
595	5	50.0	438	5	Q9NG13	Q9ng13 brachioosto	668	5	50.0	491	14	Q80181	Q80181 human immun
596	5	50.0	439	5	Q9V9N5	Q9v9n5 drosophila	669	5	50.0	491	14	Q9WLH4	Q9wlh4 human immun
597	5	50.0	440	2	Q25464	Q25464 helicobacte	670	5	50.0	491	14	Q9OEP9	Q9oep9 human immun
598	5	50.0	443	1	Q58579	Q58579 pyrococcus	671	5	50.0	491	14	Q9QBZ6	Q9qbz6 human immun
599	5	50.0	444	5	Q26861	Q26861 tryppanosoma	672	5	50.0	492	2	Q9ZND0	Q9znd0 clostridium
600	5	50.0	445	2	Q45794	Q45794 bacteroides	673	5	50.0	492	14	Q79351	Q79351 human immun
601	5	50.0	445	2	Q9K199	Q9k199 bacteroides	674	5	50.0	492	14	Q74452	Q74452 human immun
602	5	50.0	446	5	Q9N3V2	Q9n3v2 caenorhabd1	675	5	50.0	492	14	Q89291	Q89291 human immun
603	5	50.0	447	2	Q86673	Q86673 streptomyce	676	5	50.0	492	14	Q89958	Q89958 human immun

385	5	50.0	230	14	0910T5	0910T5 human immun	458	5	50.0	255	14	010731	010731 human immun
386	5	50.0	231	14	0910M3	0910M3 human immun	459	5	50.0	255	14	091746	091746 human immun
387	5	50.0	231	14	0910U5	0910U5 human immun	460	5	50.0	255	14	091742	091742 human immun
388	5	50.0	232	14	0910P8	0910P8 human immun	461	5	50.0	255	14	0916X0	0916X0 human immun
389	5	50.0	232	14	0910U9	0910U9 human immun	462	5	50.0	255	14	0916U2	0916U2 human immun
390	5	50.0	232	14	0910U8	0910U8 human immun	463	5	50.0	256	4	09N0R3	09N0R3 homo sapien
391	5	50.0	232	14	0910T4	0910T4 human immun	464	5	50.0	256	14	091709	091709 human immun
392	5	50.0	233	14	0910V0	0910V0 human immun	465	5	50.0	256	14	0916U7	0916U7 human immun
393	5	50.0	233	14	0910V4	0910V4 human immun	466	5	50.0	256	14	0916T6	0916T6 human immun
394	5	50.0	233	14	0910U7	0910U7 human immun	467	5	50.0	257	14	0917A5	0917A5 human immun
395	5	50.0	233	14	0910P8	0910P8 human immun	468	5	50.0	257	14	0916W3	0916W3 human immun
396	5	50.0	234	14	0910P0	0910P0 human immun	469	5	50.0	257	14	0916P0	0916P0 human immun
397	5	50.0	234	14	0910V9	0910V9 human immun	470	5	50.0	258	2	09K410	09K410 streptomyce
398	5	50.0	234	14	0910V6	0910V6 human immun	471	5	50.0	258	14	0916R3	0916R3 human immun
399	5	50.0	235	14	0910Z8	0910Z8 human immun	472	5	50.0	258	14	0916P4	0916P4 human immun
400	5	50.0	235	14	0916O5	0916O5 human immun	473	5	50.0	259	1	09HMK32	09HMK32 thermoplas
401	5	50.0	236	14	0910F58	0910F58 human immun	474	5	50.0	259	14	010736	010736 human immun
402	5	50.0	237	10	09F039	09F039 oryza sativ	475	5	50.0	260	5	0116Z5	0116Z5 caenorhabd1
403	5	50.0	237	14	0910Z5	0910Z5 human immun	476	5	50.0	261	14	0916R5	0916R5 human immun
404	5	50.0	237	14	0910V5	0910V5 human immun	477	5	50.0	261	14	0916P5	0916P5 human immun
405	5	50.0	238	14	0910H5	0910H5 human immun	478	5	50.0	261	14	0916N6	0916N6 human immun
406	5	50.0	238	14	0910F57	0910F57 human immun	479	5	50.0	262	14	091736	091736 human immun
407	5	50.0	241	11	092135	092135 rattus norv	480	5	50.0	263	2	P94481	P94481 bacillus su
408	5	50.0	241	14	0910T8	0910T8 human immun	481	5	50.0	263	10	09FEL5	09FEL5 caenorhabd1
409	5	50.0	243	14	0910M7	0910M7 human immun	482	5	50.0	264	14	010730	010730 human immun
410	5	50.0	244	14	0910Z4	0910Z4 human immun	483	5	50.0	265	14	010729	010729 human immun
411	5	50.0	245	14	0910Z4	0910Z4 human immun	484	5	50.0	266	4	015795	015795 homo sapien
412	5	50.0	245	14	0910Z3	0910Z3 human immun	485	5	50.0	266	5	025732	025732 plasmodium
413	5	50.0	245	14	0910Z3	0910Z3 human immun	486	5	50.0	268	14	010735	010735 human immun
414	5	50.0	245	14	0910Z3	0910Z3 human immun	487	5	50.0	271	5	021382	021382 caenorhabd1
415	5	50.0	245	14	0910Z3	0910Z3 human immun	488	5	50.0	273	2	005936	005936 pseudomonas
416	5	50.0	246	2	09RGC4	09RGC4 mycoplasma	489	5	50.0	275	4	09N0R1	09N0R1 homo sapien
417	5	50.0	246	14	0910Z1	0910Z1 human immun	490	5	50.0	276	2	P73137	P73137 synechocyst
418	5	50.0	246	14	0910Z1	0910Z1 human immun	491	5	50.0	276	5	094215	094215 caenorhabd1
419	5	50.0	246	14	0910Z1	0910Z1 human immun	492	5	50.0	278	2	068435	068435 leptospira
420	5	50.0	246	14	0910Z1	0910Z1 human immun	493	5	50.0	278	2	068437	068437 leptospira
421	5	50.0	247	14	0910Z1	0910Z1 human immun	494	5	50.0	282	5	090221	090221 caenorhabd1
422	5	50.0	247	14	0910Z1	0910Z1 human immun	495	5	50.0	287	1	0910Z3	0910Z3 pyrococcus
423	5	50.0	247	14	0910Z1	0910Z1 human immun	496	5	50.0	288	2	0910Z3	0910Z3 neisseria m
424	5	50.0	248	14	0910Z1	0910Z1 human immun	497	5	50.0	289	2	067230	067230 aquifex aeo
425	5	50.0	249	10	065058	065058 picea maria	498	5	50.0	289	2	09KYN4	09KYN4 streptomyce
426	5	50.0	249	14	0910Z1	0910Z1 human immun	499	5	50.0	291	2	006449	006449 serratia ma
427	5	50.0	249	14	0910Z1	0910Z1 human immun	500	5	50.0	291	2	007464	007464 salmonella
428	5	50.0	250	11	09W6V4	09W6V4 rattus norv	501	5	50.0	293	2	09PDL4	09PDL4 xylella fas
429	5	50.0	250	14	0910Z1	0910Z1 human immun	502	5	50.0	293	4	014836	014836 homo sapien
430	5	50.0	250	14	0910Z1	0910Z1 human immun	503	5	50.0	293	10	049020	049020 gossypium h
431	5	50.0	251	14	0910Z1	0910Z1 human immun	504	5	50.0	296	10	022812	022812 arabidopsis
432	5	50.0	252	10	0910Z1	0910Z1 arabidopsis	505	5	50.0	297	5	09VD14	09VD14 drosophila
433	5	50.0	252	14	0910Z1	0910Z1 human immun	506	5	50.0	298	5	09V090	09V090 drosophila
434	5	50.0	252	14	0910Z1	0910Z1 human immun	507	5	50.0	305	14	0910Z1	0910Z1 human immun
435	5	50.0	252	14	0910Z1	0910Z1 human immun	508	5	50.0	305	2	09K0M1	09K0M1 bacillus ha
436	5	50.0	252	14	0910Z1	0910Z1 human immun	509	5	50.0	306	10	09S0N3	09S0N3 prunus dulc
437	5	50.0	252	14	0910Z1	0910Z1 human immun	510	5	50.0	307	2	055410	055410 synechocyst
438	5	50.0	252	14	0910Z1	0910Z1 human immun	511	5	50.0	308	2	09R0Q2	09R0Q2 delnecocyst
439	5	50.0	252	14	0910Z1	0910Z1 human immun	512	5	50.0	312	14	09WIT5	09WIT5 human immun
440	5	50.0	252	14	0910Z1	0910Z1 human immun	513	5	50.0	312	14	09WIT2	09WIT2 human immun
441	5	50.0	252	14	0910Z1	0910Z1 human immun	514	5	50.0	315	5	018121	018121 caenorhabd1
442	5	50.0	253	1	059562	059562 pyrococcus	515	5	50.0	327	2	09PAP3	09PAP3 xylella fas
443	5	50.0	253	1	09V117	09V117 pyrococcus	516	5	50.0	327	10	0910D7	0910D7 oryza sativ
444	5	50.0	253	14	0910Z1	0910Z1 human immun	517	5	50.0	327	11	09QW06	09QW06 mus musculu
445	5	50.0	253	14	0910Z1	0910Z1 human immun	518	5	50.0	329	5	09Y205	09Y205 ephydattia f
446	5	50.0	253	14	0910Z1	0910Z1 human immun	519	5	50.0	329	5	09VY11	09VY11 drosophila
447	5	50.0	253	14	0910Z1	0910Z1 human immun	520	5	50.0	330	14	072628	072628 human immun
448	5	50.0	253	14	0910Z1	0910Z1 human immun	521	5	50.0	335	2	09RSG0	09RSG0 delnecocyst
449	5	50.0	254	5	021223	021223 caenorhabd1	522	5	50.0	335	5	09UD3	09UD3 caenorhabd1
450	5	50.0	254	14	0910Z1	0910Z1 human immun	523	5	50.0	337	2	053089	053089 iactobacill
451	5	50.0	254	14	0910Z1	0910Z1 human immun	524	5	50.0	339	14	091E68	091E68 human immun
452	5	50.0	254	14	0910Z1	0910Z1 human immun	525	5	50.0	342	2	0925F8	0925F8 mycobacteri
453	5	50.0	254	14	0910Z1	0910Z1 human immun	526	5	50.0	343	2	09KRG7	09KRG7 vibrio chol
454	5	50.0	254	14	0910Z1	0910Z1 human immun	527	5	50.0	343	8	092238	092238 trichophyto
455	5	50.0	254	14	0910Z1	0910Z1 human immun	528	5	50.0	344	14	0910A5	0910A5 human immun
456	5	50.0	254	14	0910Z1	0910Z1 human immun	529	5	50.0	347	5	09VR09	09VR09 drosophila
457	5	50.0	254	14	0910Z1	0910Z1 human immun	530	5	50.0	351	10	023142	023142 arabidopsis

239	5	50.0	207	14	Q9DUW8	Q9duw8	human	immun	312	5	50.0	219	14	Q9IUK0	Q9Iuk0	human	immun
240	5	50.0	207	14	Q9DUW6	Q9duw6	human	immun	313	5	50.0	220	14	Q9IY08	Q9Iy08	human	immun
241	5	50.0	207	14	Q9DUW5	Q9duw5	human	immun	314	5	50.0	220	14	Q9IVP3	Q9Ivp3	human	immun
242	5	50.0	207	14	Q9DUW3	Q9duw3	human	immun	315	5	50.0	220	14	Q9IUY5	Q9Iuy5	human	immun
243	5	50.0	207	14	Q9DUW1	Q9duw1	human	immun	316	5	50.0	220	14	Q9IUY4	Q9Iuy4	human	immun
244	5	50.0	207	14	Q9DUV9	Q9duv9	human	immun	317	5	50.0	220	14	Q9IUR3	Q9Iur3	human	immun
245	5	50.0	207	14	Q9DUV8	Q9duv8	human	immun	318	5	50.0	220	14	Q9IDR3	Q9Idr3	human	immun
246	5	50.0	207	14	Q9DUV6	Q9duv6	human	immun	319	5	50.0	220	14	Q9IDP4	Q9Idp4	human	immun
247	5	50.0	207	14	Q9DUV5	Q9duv5	human	immun	320	5	50.0	221	2	P70813	P70813	borrelia af	
248	5	50.0	207	14	Q9DUV4	Q9duv4	human	immun	321	5	50.0	221	14	Q9Q2K7	Q9q2k7	human	immun
249	5	50.0	207	14	Q9DUV1	Q9duv1	human	immun	322	5	50.0	221	14	Q9IUX5	Q9Iux5	human	immun
250	5	50.0	207	14	Q9DUU9	Q9duu9	human	immun	323	5	50.0	221	14	Q9IUW0	Q9Iuw0	human	immun
251	5	50.0	208	2	Q9RB27	Q9rb27	pseudomonas		324	5	50.0	221	14	Q9IUW6	Q9Iuw6	human	immun
252	5	50.0	208	14	Q37393	Q37393	human	immun	325	5	50.0	222	14	Q9J7A4	Q9j7a4	human	immun
253	5	50.0	208	14	Q37394	Q37394	human	immun	326	5	50.0	222	14	Q9J6Z1	Q9j6z1	human	immun
254	5	50.0	208	14	Q9J6Z4	Q9j6z4	human	immun	327	5	50.0	222	14	Q9IVN0	Q9Ivn0	human	immun
255	5	50.0	208	14	Q9IUP8	Q9iup8	human	immun	328	5	50.0	222	14	Q9IVJ3	Q9Ivj3	human	immun
256	5	50.0	208	14	Q9IUL5	Q9iul5	human	immun	329	5	50.0	222	14	Q9IVJ2	Q9Ivj2	human	immun
257	5	50.0	208	14	Q9IUK6	Q9iuk6	human	immun	330	5	50.0	222	14	Q9IUV3	Q9Iuv3	human	immun
258	5	50.0	209	14	Q9J7A6	Q9j7a6	human	immun	331	5	50.0	223	14	Q9ID08	Q9Id08	human	immun
259	5	50.0	209	14	Q9IUS7	Q9ius7	human	immun	332	5	50.0	222	14	Q9IDP0	Q9Idp0	human	immun
260	5	50.0	209	14	Q9IUR7	Q9iur7	human	immun	333	5	50.0	223	14	Q9U730	Q9u730	human	immun
261	5	50.0	209	14	Q9IUQ6	Q9iuq6	human	immun	334	5	50.0	223	14	Q9IVQ9	Q9Ivq9	human	immun
262	5	50.0	209	14	Q9IUL7	Q9iul7	human	immun	335	5	50.0	223	14	Q9IUW2	Q9Iuw2	human	immun
263	5	50.0	209	14	Q9IUK7	Q9iuk7	human	immun	336	5	50.0	223	14	Q9IUU0	Q9Iuu0	human	immun
264	5	50.0	210	14	Q9J776	Q9j776	human	immun	337	5	50.0	223	14	Q9IUR7	Q9Iur7	human	immun
265	5	50.0	210	14	Q9J6T4	Q9j6t4	human	immun	338	5	50.0	224	14	Q9J792	Q9j792	human	immun
266	5	50.0	210	14	Q9IUX8	Q9iux8	human	immun	339	5	50.0	224	14	Q9IVN9	Q9Ivn9	human	immun
267	5	50.0	210	14	Q9IUX1	Q9iux1	human	immun	340	5	50.0	224	14	Q9IVL4	Q9Ivl4	human	immun
268	5	50.0	210	14	Q9IUR4	Q9iur4	human	immun	341	5	50.0	224	14	Q9IVJ9	Q9Ivj9	human	immun
269	5	50.0	210	14	Q9DUW9	Q9duw9	human	immun	342	5	50.0	224	14	Q9IVJ8	Q9Ivj8	human	immun
270	5	50.0	211	5	Q9XU58	Q9xu58	caecorhabdi		343	5	50.0	224	14	Q9IUV8	Q9Iuv8	human	immun
271	5	50.0	211	14	Q56887	Q56887	human	immun	344	5	50.0	224	14	Q9IUY3	Q9Iuy3	human	immun
272	5	50.0	211	14	Q56888	Q56888	human	immun	345	5	50.0	224	14	Q9IUY9	Q9Iuy9	human	immun
273	5	50.0	211	14	Q56889	Q56889	human	immun	346	5	50.0	224	14	Q9IUU2	Q9Iuu2	human	immun
274	5	50.0	211	14	Q56890	Q56890	human	immun	347	5	50.0	224	14	Q9IDN9	Q9Idn9	human	immun
275	5	50.0	211	14	Q56891	Q56891	human	immun	348	5	50.0	225	14	Q9J6Z9	Q9j6z9	human	immun
276	5	50.0	211	14	Q56892	Q56892	human	immun	349	5	50.0	225	14	Q9J6R4	Q9j6r4	human	immun
277	5	50.0	211	14	Q93157	Q93157	human	immun	350	5	50.0	225	14	Q9IVR2	Q9Ivr2	human	immun
278	5	50.0	211	14	Q9IYQ3	Q9iyq3	human	immun	351	5	50.0	225	14	Q9IUX9	Q9Iux9	human	immun
279	5	50.0	211	14	Q9IVK4	Q9ivk4	human	immun	352	5	50.0	225	14	Q9IUX4	Q9Iux4	human	immun
280	5	50.0	211	14	Q9IUR0	Q9iur0	human	immun	353	5	50.0	225	14	Q9IUV7	Q9Iuv7	human	immun
281	5	50.0	211	14	Q9IUN5	Q9iun5	human	immun	354	5	50.0	225	14	Q9IDQ5	Q9Idq5	human	immun
282	5	50.0	212	14	Q9IVQ1	Q9ivq1	human	immun	355	5	50.0	226	14	Q9IVM5	Q9Ivm5	human	immun
283	5	50.0	213	14	Q9WMM0	Q9wmm0	human	immun	356	5	50.0	226	14	Q9IVL5	Q9Ivl5	human	immun
284	5	50.0	214	14	Q9J7B0	Q9j7b0	human	immun	357	5	50.0	226	14	Q9IVK7	Q9Ivk7	human	immun
285	5	50.0	214	14	Q9J7A9	Q9j7a9	human	immun	358	5	50.0	226	14	Q9IVK2	Q9Ivk2	human	immun
286	5	50.0	214	14	Q9J7A1	Q9j7a1	human	immun	359	5	50.0	226	14	Q9IVK1	Q9Ivk1	human	immun
287	5	50.0	214	14	Q9J777	Q9j777	human	immun	360	5	50.0	226	14	Q9IUV5	Q9Iuv5	human	immun
288	5	50.0	214	14	Q9IVR3	Q9ivr3	human	immun	361	5	50.0	226	14	Q9IDP7	Q9Idp7	human	immun
289	5	50.0	215	14	Q9J7B2	Q9j7b2	human	immun	362	5	50.0	226	14	Q9IDP5	Q9Idp5	human	immun
290	5	50.0	215	14	Q9IUS0	Q9ius0	human	immun	363	5	50.0	226	14	Q9IDP1	Q9Idp1	human	immun
291	5	50.0	216	14	Q9J759	Q9j759	human	immun	364	5	50.0	227	2	Q31975	Q31975	bacillus su	
292	5	50.0	216	14	Q9IUK1	Q9iuk1	human	immun	365	5	50.0	227	9	Q37974	Q37974	baacterioph	
293	5	50.0	217	14	Q9IVP9	Q9ivp9	human	immun	366	5	50.0	227	9	Q38576	Q38576	unidentifile	
294	5	50.0	217	14	Q9IUY0	Q9iuy0	human	immun	367	5	50.0	227	10	Q38764	Q38764	zea mays (m	
295	5	50.0	217	14	Q9IUY2	Q9iuy2	human	immun	368	5	50.0	227	14	Q9IVP5	Q9Ivp5	human	immun
296	5	50.0	217	14	Q9IUN8	Q9iun8	human	immun	369	5	50.0	227	14	Q9IUV8	Q9Iuv8	human	immun
297	5	50.0	217	14	Q9IUN3	Q9iun3	human	immun	370	5	50.0	227	14	Q9IUV6	Q9Iuv6	human	immun
298	5	50.0	218	14	Q9J778	Q9j778	human	immun	371	5	50.0	227	14	Q9IDQ3	Q9Idq3	human	immun
299	5	50.0	218	14	Q9J754	Q9j754	human	immun	372	5	50.0	228	14	Q9QNX0	Q9qnX0	human	immun
300	5	50.0	218	14	Q9IVP1	Q9ivp1	human	immun	373	5	50.0	228	14	Q9IVR4	Q9Ivr4	human	immun
301	5	50.0	218	14	Q9IUX2	Q9iux2	human	immun	374	5	50.0	228	14	Q9IVP2	Q9Ivp2	human	immun
302	5	50.0	218	14	Q9IDR5	Q9idr5	human	immun	375	5	50.0	228	14	Q9IVK6	Q9Ivk6	human	immun
303	5	50.0	219	4	Q9I656	Q9i656	homo sapien		376	5	50.0	228	14	Q9IUV1	Q9Iuv1	human	immun
304	5	50.0	219	14	Q9J6V9	Q9j6v9	human	immun	377	5	50.0	228	14	Q9IUV0	Q9Iuv0	human	immun
305	5	50.0	219	14	Q9IVN7	Q9ivn7	human	immun	378	5	50.0	228	14	Q9IUV3	Q9Iuv3	human	immun
306	5	50.0	219	14	Q9IVN3	Q9ivn3	human	immun	379	5	50.0	229	14	Q9J7I3	Q9j7i3	human	immun
307	5	50.0	219	14	Q9IUY2	Q9iuy2	human	immun	380	5	50.0	229	14	Q9IUV2	Q9Iuv2	human	immun
308	5	50.0	219	14	Q9IUX0	Q9iux0	human	immun	381	5	50.0	229	14	Q9IUV9	Q9Iuv9	human	immun
309	5	50.0	219	14	Q9IUV8	Q9iuv8	human	immun	382	5	50.0	230	14	Q9IVN5	Q9Ivn5	human	immun
310	5	50.0	219	14	Q9IUV5	Q9iuv5	human	immun	383	5	50.0	230	14	Q9IVJ5	Q9Ivj5	human	immun
311	5	50.0	219	14	Q9IUV4	Q9iuv4	human	immun	384	5	50.0	230	14	Q9IUV1	Q9Iuv1	human	immun

93	5	50.0	156	6	Q9N146	Q9n146 macaca mula	166	5	50.0	203	14	Q91up7	Q91up7 human	immun
94	5	50.0	157	5	Q9GxJ6	Q9gyJ6 caenorhabdi	167	5	50.0	203	14	Q91UK9	Q91uk9 human	immun
95	5	50.0	161	2	Q9FBC1	Q9fbc1 streptococc	168	5	50.0	204	14	Q9Q2J1	Q9q2j1 human	immun
96	5	50.0	161	10	Q9SSS3	Q9sss3 arabidopsis	169	5	50.0	204	14	Q9J7A8	Q9j7a8 human	immun
97	5	50.0	161	14	Q9J771	Q9j771 human	170	5	50.0	204	14	Q9J797	Q9j797 human	immun
98	5	50.0	161	14	Q9J6U4	Q9j6u4 human	171	5	50.0	204	14	Q9J782	Q9j782 human	immun
99	5	50.0	162	14	Q9J6U0	Q9j6u0 human	172	5	50.0	204	14	Q9J741	Q9j741 human	immun
100	5	50.0	163	5	Q9N3V0	Q9n3v0 caenorhabdi	173	5	50.0	204	14	Q9J6Q4	Q9j6q4 human	immun
101	5	50.0	165	14	Q9J788	Q9j788 human	174	5	50.0	204	14	Q9J6Q3	Q9j6q3 human	immun
102	5	50.0	165	14	Q9J752	Q9j752 human	175	5	50.0	204	14	Q9J6Q7	Q9j6q7 human	immun
103	5	50.0	166	4	Q9NRT0	Q9nrt0 homo sapien	176	5	50.0	205	14	Q9J1U4	Q9j1u4 human	immun
104	5	50.0	166	14	Q9J7A7	Q9j7a7 human	177	5	50.0	205	14	Q9J6P1	Q9j6p1 human	immun
105	5	50.0	166	14	Q9J6V7	Q9j6v7 human	178	5	50.0	205	14	Q9JUP5	Q9jup5 human	immun
106	5	50.0	167	14	Q9J6U3	Q9j6u3 human	179	5	50.0	205	14	Q9J1UP0	Q9j1up0 human	immun
107	5	50.0	171	5	Q9N927	Q9n927 trypanosoma	180	5	50.0	205	14	Q9J1UN6	Q9j1un6 human	immun
108	5	50.0	172	14	Q9J784	Q9j784 human	181	5	50.0	205	14	Q9J1UN0	Q9j1un0 human	immun
109	5	50.0	172	14	Q9J6S9	Q9j6s9 human	182	5	50.0	205	14	Q9J1UM2	Q9j1um2 human	immun
110	5	50.0	173	14	Q9J781	Q9j781 human	183	5	50.0	205	14	Q9J1UM0	Q9j1um0 human	immun
111	5	50.0	173	14	Q9J6Z3	Q9j6z3 human	184	5	50.0	206	14	Q9J772	Q9j772 human	immun
112	5	50.0	175	14	Q9J6W9	Q9j6w9 human	185	5	50.0	206	14	Q9J6W1	Q9j6w1 human	immun
113	5	50.0	177	14	Q9J791	Q9j791 human	186	5	50.0	206	14	Q9J6V6	Q9j6v6 human	immun
114	5	50.0	178	14	Q9J751	Q9j751 human	187	5	50.0	206	14	Q9J1UK3	Q9j1uk3 human	immun
115	5	50.0	178	14	Q9J750	Q9j750 human	188	5	50.0	207	14	Q9JWW6	Q9jww6 human	immun
116	5	50.0	179	14	Q9J747	Q9j747 human	189	5	50.0	207	14	Q9JWW4	Q9jww4 human	immun
117	5	50.0	179	10	Q9ZQD4	Q9zqd4 arabidopsis	190	5	50.0	207	14	Q9JWW3	Q9jww3 human	immun
118	5	50.0	179	14	Q9J786	Q9j786 human	191	5	50.0	207	14	Q9JWW2	Q9jww2 human	immun
119	5	50.0	179	14	Q9J6V2	Q9j6v2 human	192	5	50.0	207	14	Q9JWW1	Q9jww1 human	immun
120	5	50.0	180	14	Q9J737	Q9j737 human	193	5	50.0	207	14	Q9JWW9	Q9jww9 human	immun
121	5	50.0	181	14	Q9J7A3	Q9j7a3 human	194	5	50.0	207	14	Q9JWW8	Q9jww8 human	immun
122	5	50.0	181	14	Q9J793	Q9j793 human	195	5	50.0	207	14	Q9QNX4	Q9qnx4 human	immun
123	5	50.0	181	14	Q9J780	Q9j780 human	196	5	50.0	207	14	Q9Q2L6	Q9q2l6 human	immun
124	5	50.0	183	14	Q9J706	Q9j706 human	197	5	50.0	207	14	Q9Q2L5	Q9q2l5 human	immun
125	5	50.0	184	14	Q9J6T7	Q9j6t7 human	198	5	50.0	207	14	Q9Q2K8	Q9q2k8 human	immun
126	5	50.0	185	10	Q9MAI5	Q9mai5 arabidopsis	199	5	50.0	207	14	Q9Q2K5	Q9q2k5 human	immun
127	5	50.0	186	14	Q9WKB9	Q9wkb9 human	200	5	50.0	207	14	Q9Q2K3	Q9q2k3 human	immun
128	5	50.0	186	14	Q9J6Z5	Q9j6z5 human	201	5	50.0	207	14	Q9Q2K2	Q9q2k2 human	immun
129	5	50.0	187	5	Q9N344	Q9n344 caenorhabdi	202	5	50.0	207	14	Q9Q2J9	Q9q2j9 human	immun
130	5	50.0	187	14	Q9J1UP3	Q9j1up3 human	203	5	50.0	207	14	Q9Q2J9	Q9q2j9 human	immun
131	5	50.0	190	1	Q9V1G6	Q9v1g6 pyrococcus	204	5	50.0	207	14	Q9Q2J9	Q9q2j9 human	immun
132	5	50.0	190	14	Q9J3038	Q9j3038 human	205	5	50.0	207	14	Q9Q2J7	Q9q2j7 human	immun
133	5	50.0	190	14	Q9J6N5	Q9j6n5 human	206	5	50.0	207	14	Q9Q2J6	Q9q2j6 human	immun
134	5	50.0	192	14	Q9J390	Q9j390 hepatitis c	207	5	50.0	207	14	Q9Q2J5	Q9q2j5 human	immun
135	5	50.0	192	14	Q9J733	Q9j733 human	208	5	50.0	207	14	Q9Q2J4	Q9q2j4 human	immun
136	5	50.0	192	14	Q9J6N9	Q9j6n9 human	209	5	50.0	207	14	Q9Q2J3	Q9q2j3 human	immun
137	5	50.0	193	14	Q9J7B3	Q9j7b3 human	210	5	50.0	207	14	Q9Q2J2	Q9q2j2 human	immun
138	5	50.0	195	14	Q9J769	Q9j769 human	211	5	50.0	207	14	Q9Q2J2	Q9q2j2 human	immun
139	5	50.0	195	14	Q9J731	Q9j731 human	212	5	50.0	207	14	Q9Q2I6	Q9q2i6 human	immun
140	5	50.0	195	14	Q9J6U1	Q9j6u1 human	213	5	50.0	207	14	Q9Q2I5	Q9q2i5 human	immun
141	5	50.0	195	14	Q9J1UQ4	Q9j1uq4 human	214	5	50.0	207	14	Q9Q2I3	Q9q2i3 human	immun
142	5	50.0	196	14	Q9J796	Q9j796 human	215	5	50.0	207	14	Q9Q2I2	Q9q2i2 human	immun
143	5	50.0	197	14	Q9J1UN4	Q9j1un4 human	216	5	50.0	207	14	Q9J1US6	Q9j1us6 human	immun
144	5	50.0	198	14	Q9J787	Q9j787 human	217	5	50.0	207	14	Q9J1UN2	Q9j1un2 human	immun
145	5	50.0	198	14	Q9J1UL8	Q9j1ul8 human	218	5	50.0	207	14	Q9J1UN1	Q9j1un1 human	immun
146	5	50.0	198	14	Q9J1UL3	Q9j1ul3 human	219	5	50.0	207	14	Q9J1UN1	Q9j1un1 human	immun
147	5	50.0	198	14	Q9J1UL2	Q9j1ul2 human	220	5	50.0	207	14	Q9J1U08	Q9j1u08 human	immun
148	5	50.0	199	14	Q9J6Z7	Q9j6z7 human	221	5	50.0	207	14	Q9J1U05	Q9j1u05 human	immun
149	5	50.0	199	14	Q9J1UP4	Q9j1up4 human	222	5	50.0	207	14	Q9J1U03	Q9j1u03 human	immun
150	5	50.0	199	14	Q9J1UM8	Q9j1um8 human	223	5	50.0	207	14	Q9J1U01	Q9j1u01 human	immun
151	5	50.0	200	14	Q9J6Q2	Q9j6q2 human	224	5	50.0	207	14	Q9J1U00	Q9j1u00 human	immun
152	5	50.0	200	14	Q9J1UP2	Q9j1up2 human	225	5	50.0	207	14	Q9J1U29	Q9j1u29 human	immun
153	5	50.0	201	14	Q9J744	Q9j744 human	226	5	50.0	207	14	Q9J1U25	Q9j1u25 human	immun
154	5	50.0	201	14	Q9J1VR7	Q9j1vr7 human	227	5	50.0	207	14	Q9J1U24	Q9j1u24 human	immun
155	5	50.0	201	14	Q9J1US1	Q9j1us1 human	228	5	50.0	207	14	Q9J1U21	Q9j1u21 human	immun
156	5	50.0	201	14	Q9J1UL1	Q9j1ul1 human	229	5	50.0	207	14	Q9J1U20	Q9j1u20 human	immun
157	5	50.0	202	14	Q9J7A2	Q9j7a2 human	230	5	50.0	207	14	Q9J1U19	Q9j1u19 human	immun
158	5	50.0	202	14	Q9J1UT1	Q9j1ut1 human	231	5	50.0	207	14	Q9J1U13	Q9j1u13 human	immun
159	5	50.0	202	14	Q9J1UP6	Q9j1up6 human	232	5	50.0	207	14	Q9J1U12	Q9j1u12 human	immun
160	5	50.0	202	14	Q9J1UM4	Q9j1um4 human	233	5	50.0	207	14	Q9J1U11	Q9j1u11 human	immun
161	5	50.0	202	14	Q9J1UL9	Q9j1ul9 human	234	5	50.0	207	14	Q9J1U09	Q9j1u09 human	immun
162	5	50.0	202	14	Q9J1UL0	Q9j1ul0 human	235	5	50.0	207	14	Q9J1U09	Q9j1u09 human	immun
163	5	50.0	203	14	Q9J6T3	Q9j6t3 human	236	5	50.0	207	14	Q9J1U07	Q9j1u07 human	immun
164	5	50.0	203	14	Q9J1US5	Q9j1us5 human	237	5	50.0	207	14	Q9J1U04	Q9j1u04 human	immun
165	5	50.0	203	14	Q9J1UQ0	Q9j1uq0 human	238	5	50.0	207	14	Q9J1U03	Q9j1u03 human	immun

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:42:49 ; Search time 58.01 Seconds  
(without alignments)  
22.807 Million cell updates/sec

Title: US-09-372-036-31  
Perfect score: 10  
Sequence: 1 PVAPEQEVKK 10

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL\_16: \*  
1: sp\_archaea: \*  
2: sp\_bacteria: \*  
3: sp\_fungi: \*  
4: sp\_human: \*  
5: sp\_invertebrate: \*  
6: sp\_mammal: \*  
7: sp\_mhc: \*  
8: sp\_organelle: \*  
9: sp\_phage: \*  
10: sp\_plant: \*  
11: sp\_rodent: \*  
12: sp\_unclassified: \*  
13: sp\_vertebrate: \*  
14: sp\_virus: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	100.0	86	2	Q9RNU3
2	10	100.0	478	2	Q9RNU3
3	7	70.0	94	14	067673
4	7	70.0	227	2	Q9PM45
5	6	60.0	102	2	Q9HYT7
6	6	60.0	116	2	Q9X9T5
7	6	60.0	174	14	Q930A1
8	6	60.0	274	2	Q9KXZ4
9	6	60.0	486	5	Q26617
10	6	60.0	609	4	Q9UNU0
11	6	60.0	624	4	Q9NZM6
12	6	60.0	748	10	Q22907
13	6	60.0	819	5	Q9YV48
14	6	60.0	988	14	Q98830
15	6	60.0	988	14	Q870A1
16	6	60.0	989	14	P90288
17	6	60.0	1230	2	Q25772
18	6	60.0	1653	2	Q9L8G3
19	5	50.0	23	11	Q63334

20	5	50.0	52	6	Q9GL38	Q9G138 bos taurus
21	5	50.0	75	2	Q9RRH1	Q9RRH1 deinococcus
22	5	50.0	77	14	Q91692	Q91692 human immun
23	5	50.0	82	5	Q17728	Q17728 caenorhabdi
24	5	50.0	88	14	Q9DR28	Q9DR28 human immun
25	5	50.0	91	14	Q9WJ74	Q9WJ74 human immun
26	5	50.0	95	14	Q68302	Q68302 hepatitis c
27	5	50.0	102	14	Q68301	Q68301 hepatitis c
28	5	50.0	108	14	Q9IE25	Q9IE25 human immun
29	5	50.0	112	4	Q9UG25	Q9UG25 homo sapien
30	5	50.0	113	14	Q93051	Q93051 human immun
31	5	50.0	113	14	Q93052	Q93052 human immun
32	5	50.0	113	14	Q93053	Q93053 human immun
33	5	50.0	113	14	Q9WKA3	Q9WKA3 human immun
34	5	50.0	113	14	Q9WKA2	Q9WKA2 human immun
35	5	50.0	113	14	Q9WKA1	Q9WKA1 human immun
36	5	50.0	113	14	Q9WKA8	Q9WKA8 human immun
37	5	50.0	113	14	Q9WKA6	Q9WKA6 human immun
38	5	50.0	113	14	Q9WKA9	Q9WKA9 human immun
39	5	50.0	113	14	Q9WKA3	Q9WKA3 human immun
40	5	50.0	113	14	Q9WKA9	Q9WKA9 human immun
41	5	50.0	113	14	Q9WKA9	Q9WKA9 human immun
42	5	50.0	113	14	Q9WKA8	Q9WKA8 human immun
43	5	50.0	113	14	Q9WKA8	Q9WKA8 human immun
44	5	50.0	113	14	Q9WKA8	Q9WKA8 human immun
45	5	50.0	113	14	Q9WKA7	Q9WKA7 human immun
46	5	50.0	113	14	Q9WKA5	Q9WKA5 human immun
47	5	50.0	114	10	Q9SVK9	Q9SVK9 arabidopsis
48	5	50.0	121	5	Q9K3V1	Q9K3V1 caenorhabdi
49	5	50.0	122	5	Q9VY28	Q9VY28 drosophila
50	5	50.0	124	2	Q84412	Q84412 chlamydia t
51	5	50.0	124	2	Q9Z818	Q9Z818 chlamydia p
52	5	50.0	124	5	Q9K045	Q9K045 caenorhabdi
53	5	50.0	127	2	Q9K6R1	Q9K6R1 yersinia pe
54	5	50.0	130	14	Q560B5	Q560B5 human immun
55	5	50.0	130	14	Q9Q145	Q9Q145 human immun
56	5	50.0	130	14	Q9Q142	Q9Q142 human immun
57	5	50.0	130	14	Q9Q141	Q9Q141 human immun
58	5	50.0	130	14	Q9Q140	Q9Q140 human immun
59	5	50.0	130	14	Q9Q135	Q9Q135 human immun
60	5	50.0	130	14	Q9Q125	Q9Q125 human immun
61	5	50.0	130	14	Q9QK25	Q9QK25 human immun
62	5	50.0	130	14	Q9QK26	Q9QK26 human immun
63	5	50.0	130	14	Q9QK21	Q9QK21 human immun
64	5	50.0	130	14	Q9QKX0	Q9QKX0 human immun
65	5	50.0	130	14	Q9QKX8	Q9QKX8 human immun
66	5	50.0	130	14	Q9QKX5	Q9QKX5 human immun
67	5	50.0	130	14	Q9QKX1	Q9QKX1 human immun
68	5	50.0	131	14	Q9Q144	Q9Q144 human immun
69	5	50.0	131	14	Q9Q118	Q9Q118 human immun
70	5	50.0	131	14	Q9Q116	Q9Q116 human immun
71	5	50.0	131	14	Q9Q108	Q9Q108 human immun
72	5	50.0	131	14	Q9QKX9	Q9QKX9 human immun
73	5	50.0	131	14	Q9QKX7	Q9QKX7 human immun
74	5	50.0	131	14	Q9QKX8	Q9QKX8 human immun
75	5	50.0	131	14	Q9QKX6	Q9QKX6 human immun
76	5	50.0	132	2	Q9K2C0	Q9K2C0 chlamydia p
77	5	50.0	132	14	Q9Q123	Q9Q123 human immun
78	5	50.0	133	2	Q87645	Q87645 methylococ
79	5	50.0	133	3	Q9UTJ5	Q9UTJ5 schizosacch
80	5	50.0	133	14	Q9QK05	Q9QK05 human immun
81	5	50.0	133	14	Q9QKX5	Q9QKX5 human immun
82	5	50.0	134	2	Q9P1Y9	Q9P1Y9 chlamydia m
83	5	50.0	137	5	Q9K3U4	Q9K3U4 caenorhabdi
84	5	50.0	140	5	Q9K3B1	Q9K3B1 caenorhabdi
85	5	50.0	141	1	Q9YAL8	Q9YAL8 aeropyrum p
86	5	50.0	141	1	Q05690	Q05690 saccharomyc
87	5	50.0	142	14	Q9J6S4	Q9J6S4 human immun
88	5	50.0	143	2	Q07266	Q07266 mycobacteri
89	5	50.0	147	2	Q9WYK3	Q9WYK3 thermotoga
90	5	50.0	147	5	Q76560	Q76560 caenorhabdi
91	5	50.0	148	10	Q9FPI3	Q9FPI3 arabidopsis
92	5	50.0	149	14	Q68732	Q68732 hepatitis c
93	5	50.0	152	14	Q9J756	Q9J756 human immun

**This Page Blank (uspto)**

Wed Aug 15 13:35:24 2001

us-09-372-036-31.011.rpr

Page 19

Search completed: August 15, 2001, 12:34:46  
Job time: 111 sec

---



A:Residues: 1-211 <MIL>  
A:Cross-references: EMBL:283226; PIDN:CAB05723.1; GSPDB:GN00023; CESP:F43D2.3  
A:Experimental source: clone F43D2  
C:Genetics:  
A:Gene: CESP:F43D2.3  
A:Map position: 5  
A:introns: 44/3; 126/3; 166/3

Query Match 50.0%; Score 5; DB 2; Length 211;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
|||||  
DB 146 PVAPT 150

RESULT 47  
late 33k protein - human adenovirus 41  
N:Alternate names: 33k phosphoprotein  
C:Species: Mastadenovirus h41 (human adenovirus 41)  
C>Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 26-Aug-1999  
C:Accession: S10212; S04851  
R:Slamende, S.B.; Plenzl, N.J.; Velarde Jr., J.; Plenzl, D.; Luftig, R.B.  
Nucleic Acids Res. 18, 3069, 1990  
A:Title: Nucleotide sequence of the region coding for 100k and 33k proteins of human ent  
A:Reference number: S10206; MID:90272433  
A:Accession: S10212  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-217 <SLF>  
A:Cross-references: EMBL:X52532  
R:Plenzl, N.; Velarde Jr., J.; Plenzl, D.; Luftig, R.B.  
Nucleic Acids Res. 17, 5398, 1989  
A:Title: Nucleotide sequence of human enteric adenovirus type 41 hexon-associated protei  
A:Reference number: S04851; MID:89345113  
A:Accession: S04851  
A:Molecule type: DNA  
A:Residues: 203-217 <PIE>  
A:Cross-references: EMBL:X15137; NID:958481; PIDN:CAA33236.1; PID:9832970  
C:Genetics:  
A:introns: 120/1  
A:Superfamily: adenovirus late 33k protein  
C:Keywords: late protein; phosphoprotein

Query Match 50.0%; Score 5; DB 2; Length 217;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
|||||  
DB 118 PVAPT 122

RESULT 48  
probable transcription factor - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T37859  
R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, February 1996  
A:Reference number: Z21750  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-217 <DEV>  
A:Cross-references: EMBL:Z69795; PIDN:CAA93690.1; GSPDB:GN00066; SPDB:SPAC1768.07  
A:Experimental source: strain 972h-; cosmid c1768  
C:Genetics:

A:Gene: SPDB:SPAC1768.07  
A:Map position: 1

Query Match 50.0%; Score 5; DB 2; Length 217;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 QEVK 10  
|||||  
DB 191 QEVK 195

RESULT 49  
transcription regulator Tena [imported] - Lactococcus lactis subsp. lactis (strain IL  
C6847  
C:Species: Lactococcus lactis subsp. lactis  
C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C86847  
R:Polotin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Eh  
Genome Res. in press, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium.  
A:Reference number: A86625  
A:Accession: C86847  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-218 <STO>  
A:Cross-references: GB:AE005176; NID:q12724803; PIDN:AAK05877.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: tena  
C:Superfamily: transcription activator tena

Query Match 50.0%; Score 5; DB 2; Length 218;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
|||||  
DB 104 PVAPT 108

RESULT 50  
protein serine/threonine phosphatase [imported] - Lactococcus lactis subsp. lactis (s  
H86860  
C:Species: Lactococcus lactis subsp. lactis  
C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: H86860  
R:Polotin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Eh  
Genome Res. in press, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium.  
A:Reference number: A86625  
A:Accession: H86860  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-221 <STO>  
A:Cross-references: GB:AE005176; NID:q12724922; PIDN:AAK05986.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: pppl  
C:Superfamily: conserved hypothetical protein ylo0; conserved hypothetical protein y1

Query Match 50.0%; Score 5; DB 2; Length 221;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PROEV 8  
|||||  
DB 216 PROEV 220

ribosomal protein L29 - Mycoplasma genitalium  
 C:Species: Mycoplasma genitalium  
 C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 07-Dec-1999  
 C:Accession: F64217  
 R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;  
 M.; Fuhmann, J.; Nguyen, D.; Uitterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.  
 C.A.; Venter, J.C.  
 Science 270, 397-403, 1995  
 A:Title: The minimal gene complement of Mycoplasma genitalium.  
 A:Reference number: A64200; MUID:96026346  
 A:Accession: F64217  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-200 <TIGR>  
 A:Cross-references: GB:U39695; GB:L43967; NID:g1045833; PID:g1045843; TIGR:MG159  
 A:Experimental source: strain G-37  
 A:Genetics:  
 A:Genetic code: SGC3

Query Match 50.0%; Score 5; DB 2; Length 200;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 OEVRK 10  
 |||||  
 DB 170 OEVRK 174

RESULT 43  
 IMPSB  
 Immunity protein - phage SP-beta  
 C:Species: phage SP-beta  
 A:Note: host Bacillus subtilis  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 31-Oct-1997  
 C:Accession: A24499  
 R:McLaughlin, J.R.; Wong, H.C.; Ting, Y.E.; Van Arsdel, J.N.; Chang, S.  
 J. Bacteriol. 167, 952-959, 1986  
 A:Title: Control of lysogeny and immunity of Bacillus subtilis temperate bacteriophage S  
 A:Reference number: A24499; MUID:86304188  
 A:Accession: A24499  
 A:Molecule type: DNA  
 A:Residues: 1-201 <MCU>  
 A:Cross-references: GB:M13821  
 C:Genetics:  
 A:Gene: d  
 C:Superfamily: phage SP-beta immunity protein  
 C:Keywords: Immunity protein

Query Match 50.0%; Score 5; DB 1; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PTQEV 8  
 |||||  
 DB 156 PTQEV 160

RESULT 44  
 A71326  
 Probable V-type ATPase, subunit D (atpD-1) - syphilis spirochete  
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
 C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
 C:Accession: A71326  
 R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
 rson, J.; Kralik, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uitterback, T.; MDC  
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998  
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
 A:Reference number: A71326; MUID:98332770  
 A:Accession: A71326  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
 A:Residues: 1-206 <COI>  
 A:Cross-references: GB:AE001220; GB:AE000520; NID:g33322705; PIDN:AA65414.1; PID:g3332  
 A:Experimental source: strain Nichols  
 C:Genetics:  
 A:Gene: TP0428  
 C:Superfamily: H+-transporting ATPase chain D

Query Match 50.0%; Score 5; DB 2; Length 206;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 OEVRK 10  
 |||||  
 DB 80 OEVRK 84

RESULT 45  
 S04827  
 gene 40A protein - phage phi-80  
 N:Contains: gene 40A protein; gene 40B protein  
 C:Species: phage phi-80  
 C:Date: 19-May-1994 #sequence\_revision 24-May-1996 #text\_change 04-Mar-2000  
 C:Accession: S04827; S04826  
 R:Ogawa, T.; Ogawa, H.; Tomizawa, J.I.  
 J. Mol. Biol. 202, 537-550, 1988  
 A:Title: Organization of the early region of bacteriophage phi-80. Genes and proteins  
 A:Reference number: S01776; MUID:89011978  
 A:Accession: S04827  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-209 <OGA>  
 A:Cross-references: EMBL:X13065; NID:g14800; PIDN:CAA31470.1; PID:g14803  
 A:Genetics: G40A  
 A:Accession: S04826  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 115-209 <OG2>  
 A:Cross-references: EMBL:X13065; NID:g14800; PIDN:CAA31469.1; PID:g14802  
 A:Genetics: G40B  
 A:Note: the authors show Met-115 as alternative initiator  
 C:Genetics: <G40A>  
 A:Gene: 40A  
 C:Genetics: <G40B>  
 A:Gene: 40B  
 C:Superfamily: phage phi-80 gene 40A protein  
 C:Keywords: alternative initiators  
 F:1-209/Product: gene 40A protein #status predicted <MAT1>  
 F:115-209/Product: gene 40B protein #status predicted <MAT2>

Query Match 50.0%; Score 5; DB 2; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TOEVR 9  
 |||||  
 DB 59 TOEVR 63

RESULT 46  
 T22122  
 hypothetical protein F43D2.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T22122  
 R:Mortimore, B.  
 submitted to the EMBL Data Library, December 1996  
 A:Reference number: T22122  
 A:Accession: T22122  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA

A:Gene: F12M16.6  
A:Map position: 1

## Query Match

50.0%; Score 5; DB 2; Length 165;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
|||||  
DB 12 PVAPT 16

## RESULT 38

JC4806  
core protein G - phage phi-K

C:Species: phage phi-K  
C:Date: 15-Aug-1996 #sequence\_revision 15-Oct-1996 #text\_change 20-Jun-2000

C:Accession: JC4806; B04253; A04253

R:Kodaira, K.; Oki, M.; Kakikawa, M.; Kimoto, H.; Taketo, A.

J. Biochem. 119, 1062-1069, 1996

A:Title: The virion proteins encoded by bacteriophage phi-K and its host-range mutant ph

A:Reference number: JC4804; MUID:96424987

A:Accession: JC4806

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-187 <KOD>

A:Cross-references: EMBL:X60323; NID:g1478118; PIDN:CAA42892.1; PID:g1478127

R:Stms, J.; Capon, D.; Dressler, D.

J. Biol. Chem. 254, 12615-12628, 1979

A:Title: dnaG (primase)-dependent origins of DNA replication. Nucleotide sequences of th

A:Reference number: A92247; MUID:80049950

A:Accession: B04253

A:Molecule type: DNA

A:Residues: 165-187 <SIM>

C:Comment: This protein is one of the structural components of the bacteriophage capsid.

C:Genetics:

A:Gene: G

C:Superfamily: phage phi-X174 gene G protein

C:Keywords: capsid protein

Query Match 50.0%; Score 5; DB 2; Length 187;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
|||||  
DB 26 PVAPT 30

## RESULT 39

H75162  
anthranilate synthase component II (trpG) PAB2046 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: H75162

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: H75162

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-190 <RAW>

A:Cross-references: GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB49383.1; PID:g545789

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB2046

C:Superfamily: glutamine amidotransferase; trpG homology  
F:2-186/Domain: trpG homology <TRG>

Query Match 50.0%; Score 5; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 QEVKK 10  
|||||  
DB 35 QEVKK 39

## RESULT 40

A41050  
repressor protein c - phage Mu

C:Species: phage Mu

C:Date: 03-Apr-1992 #sequence\_revision 03-Apr-1992 #text\_change 14-Jul-1994

C:Accession: A41050; A41049

R:Geskens, V.; Vogel, J.L.; Grimaud, R.; Desmet, L.; Higgins, N.P.; Toussaint, A.

J. Bacteriol. 173, 6578-6585, 1991

A:Title: Frameshift mutations in the bacteriophage Mu repressor gene can confer a tra

A:Reference number: A41050; MUID:92011434

A:Accession: A41050

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-196 <SEU>

R:Vogel, J.L.; Li, Z.J.; Howe, M.M.; Toussaint, A.; Higgins, N.P.

J. Bacteriol. 173, 6568-6577, 1991

A:Title: Temperature-sensitive mutations in the bacteriophage Mu c repressor locate a

A:Reference number: A41049; MUID:92011433

A:Accession: A41049

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-196 <VOG>

Query Match 50.0%; Score 5; DB 2; Length 196;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 QEVKK 10  
|||||  
DB 190 QEVKK 194

## RESULT 41

F86832  
hypothetical protein yrbB [imported] - Lactococcus lactis subsp. lactis (strain IL140

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 23-Mar-2001

C:Accession: F86832

R:Bohlochin, A.; Winkler, P.; Mauger, S.; Jalllon, O.; Malarme, K.; Weissenbach, J.; Eh

Genome Res. In press, 2001

A:Title: The complete genome sequence of the lactic acid bacterium.

A:Reference number: A86525

A:Accession: F86832

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-199 <STO>

A:Cross-references: GB:AE005176; NID:g12724674; PIDN:AAK05760.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: yrbB

Query Match 50.0%; Score 5; DB 2; Length 199;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAPRQ 6  
|||||  
DB 96 VAPRQ 100

## RESULT 42

F64217

A:Reference number: A86625  
A:Accession: D86699  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-175 <STO>  
A:Cross-References: GB:AE005176; NID:g12723488; PIDN:AAK04694.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: yf11

Query Match 50.0%; Score 5; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 QEVKK 10  
DB 94 QEVKK 98

## RESULT 34

hypothetical protein Atg37200 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: G84789  
R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.; Eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: G84789  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-179 <STO>  
A:Cross-References: GB:AE002093; NID:g4371281; PIDN:AAD18139.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: Atg37200  
A:Map position: 2

Query Match 50.0%; Score 5; DB 2; Length 179;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
DB 15 PVAPT 19

## RESULT 35

elongation factor P [imported] - Chlamydomonas reinhardtii (strain J138)  
C:Species: Chlamydomonas reinhardtii (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
C:Accession: H86513  
R:Shital, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishitani, F.; Ouchi, K.; Shiba, T.; Is Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of Chlamydomonas reinhardtii J138.  
A:Reference number: A86491; MUID:20330349  
A:Accession: H86513  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-185 <STO>  
A:Cross-References: GB:BA000008; NID:g8978557; PIDN:BA038394.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: efp\_1  
C:Superfamily: translation elongation factor EF-P

Query Match 50.0%; Score 5; DB 2; Length 185;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVK 9  
DB 57 TOEVK 61

## RESULT 36

translation elongation factor P CP0584 [imported] - Chlamydomonas reinhardtii (strain J138)  
N:Alternate names: translation elongation factor EF-P  
C:Species: Chlamydomonas reinhardtii (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
C:Accession: G72110; E81560  
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydomonas reinhardtii and C. trachomatis.  
A:Reference number: A72000; MUID:99206606  
A:Accession: G72110  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-185 <ARN>  
A:Cross-References: GB:AE001604; GB:AE001363; NID:g4376438; PIDN:AAD18337.1; PID:g437  
A:Experimental source: strain CWD029  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydomonas trachomatis Mopn and Chlamydomonas reinhardtii AR39  
A:Reference number: A81500; MUID:20150255  
A:Accession: E81560  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-185 <REA>  
A:Cross-References: GB:AE002217; GB:AE002161; NID:g7189493; PIDN:AAF38402.1; PID:g718  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: efp\_1; CP0584  
C:Superfamily: translation elongation factor EF-P

Query Match 50.0%; Score 5; DB 2; Length 185;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVK 9  
DB 57 TOEVK 61

## RESULT 37

protein F12M16.6 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: D96572  
R:Theodoridis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, aasen, N.F.; Hughes, B.; Hulzar, L. Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Malt, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H. A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W. A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: D96572  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-185 <STO>  
A:Cross-References: GB:AE005173; NID:g7769876; PIDN:AAF9554.1; GSPDB:GN00141  
C:Genetics:

A:Introns: 45/3; 58/1; 75/1; 102/3

Query Match 50.0%; Score 5; DB 2; Length 147;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APPOE 7  
|||||  
DB 53 APPOE 57

RESULT 29

STH-21 protein - potato

C:Species: Solanum tuberosum (potato)

C:Date: 13-Jan-1995 #sequence\_revision 09-May-1997 #text\_change 20-Aug-1999

C:Accession: S35162; S11869

R:Matton, D.P.; Prescott, G.; Bertrand, C.; Camlrand, A.; Brisson, N.

Plant Mol. Biol. 22, 279-291, 1993

A:Title: Identification of cis-acting elements involved in the regulation of the pathoge

A:Reference number: S35161; MUID:93283632

A:Accession: S35162

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-155 <MAT>

A:Cross-references: EMBL:M29042; NID:g169577; PIDN:AA02829.1; PID:g169578

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1990

A:Gene: STH-21

A:Introns: 57/1

C:Superfamily: pathogenesis-related protein

Query Match 50.0%; Score 5; DB 2; Length 155;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
|||||  
DB 13 PVAPT 17

RESULT 30

T25638

hypothetical protein C46H11.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T25638

R:Miller, N.; Bradshaw, H.; Wamsley, P.

submitted to the EMBL Data Library, February 1997

A:Description: The sequence of C. elegans cosmid C46H11.

A:Reference number: Z20061

A:Accession: T25638

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-157 <MTL>

A:Cross-references: EMBL:U080314; PIDN:AA042359.1; GSPDB:GN00019; CESP:C46H11.5

A:Experimental source: strain Bristol N2; clone C46H11

C:Genetics:

A:Gene: CESP:C46H11.5

A:Map position: 1

A:Introns: 12/1; 44/2

Query Match 50.0%; Score 5; DB 2; Length 157;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APPOE 7  
|||||  
DB 86 APPOE 90

RESULT 31

C96568

hypothetical protein F6D8.6 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: C96568

R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

ansen, N.E.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: C96568

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1161 <STO>

A:Cross-references: GB:AE005173; NID:g5903065; PIDN:AA05624.1; GSPDB:GN00141

C:Genetics:

A:Gene: F6D8.6

A:Map position: 1

Query Match 50.0%; Score 5; DB 2; Length 161;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
|||||  
DB 40 PVAPT 44

RESULT 32

S07291

repressor protein CI - phage Mu

C:Species: phage Mu

C:Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 08-Oct-1999

C:Accession: S07291

R:Priess, H.; Kamp, D.; Kahmann, R.; Brauer, B.; Dellus, H.

Mol. Gen. Genet. 186, 315-321, 1982

A:Title: Nucleotide sequence of the immunity region of bacteriophage Mu.

A:Reference number: S07291; MUID:83012203

A:Accession: S07291

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-174 <PRI>

A:Cross-references: EMBL:V01464; NID:g15807; PIDN:CA024711.1; PID:g15808

C:Keywords: DNA binding; transcription regulation

Query Match 50.0%; Score 5; DB 2; Length 174;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 OEYK 10  
|||||  
DB 168 OEYK 172

RESULT 33

D86699

hypothetical protein yfii [imported] - Lactococcus lactis subsp. lactis (strain IL140

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 23-Mar-2001

C:Accession: D86699

R:Boletín, A.; Wincker, P.; Mauger, S.; Jallón, O.; Malarme, K.; Weissenbach, J.; Eh

Genome Res. in press, 2001

A:Title: The complete genome sequence of the lactic acid bacterium.

Db 39 OEYKK 43

RESULT 24

T50083  
probable cytidine deaminase [imported] - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: *Schizosaccharomyces pombe*  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 21-Jul-2000  
C:Accession: T50083  
R:Saunders, D.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, November 1999  
A:Reference number: Z25035  
A:Accession: T50083  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-133 <SAU>  
A:Cross-references: EMBL:AL132984; PIDN:CAH61215.1; GSPDB:GN00066; SPDB:SPAC1556.04C  
A:Experimental source: strain 972h(-); cosmid c1556  
C:Genetics:  
A:Gene: SPDB:SPAC1556.04C  
A:Map position: 1  
A:Introns: 47/1; 67/3; 95/2; 121/3  
C:Superfamily: cytidine deaminase

Query Match 50.0%; Score 5; DB 2; Length 133;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 25

AB1676  
conserved hypothetical protein TC0687 [imported] - *Chlamydia muridarum* (strain N199)

C:Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MOPn  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Sep-2000  
R:Read, T.D.; Brundham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwin, M.; Neilson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of *Chlamydia trachomatis* MOPn and *Chlamydia pneumoniae* AR39.  
A:Reference number: AB1500; MUID:20150255  
A:Accession: AB1676  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-134 <TER>  
A:Cross-references: GB:AE002337; GB:AE002160; NID:g7190714; PIDN:AAF39504.1; PID:g719071  
A:Experimental source: strain N199 (MOPn)  
C:Genetics:  
A:Gene: TC0687  
C:Superfamily: *Escherichia coli* dosage-dependent dnaK suppressor protein dksA

Query Match 50.0%; Score 5; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 OEYKK 10  
Db 41 OEYKK 45

RESULT 26

E72580  
hypothetical protein APE1924 - *Aeropyrum pernix* (strain K1)

C:Species: *Aeropyrum pernix*  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
C:Accession: E72580  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hukawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aero*

A:Reference number: A72450; MUID:99310339  
A:Accession: E72580  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-141 <KAW>  
A:Cross-references: DDBJ:AP000062; NID:g5105244; PIDN:BAAB0930.1; PID:d1044716; PID:g  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE1924

Query Match 50.0%; Score 5; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAPTO 6  
Db 66 VAPTO 70

RESULT 27

G72384  
conserved hypothetical protein - *Thermotoga maritima* (strain MSB8)

C:Species: *Thermotoga maritima*  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: G72384  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MUID:99287316  
A:Accession: G72384  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-147 <ARN>  
A:Cross-references: GB:AE001717; GB:AE000512; NID:g4980871; PIDN:AAD35456.1; PID:g498  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM0369

Query Match 50.0%; Score 5; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 OEYKK 10  
Db 127 OEYKK 131

RESULT 28

T33327  
hypothetical protein C40D2.4 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T33327  
R:Madson, C.; Wamsley, P.; O'Brien, D.  
submitted to the EMBL Data Library, July 1998  
A:Description: The sequence of *Caenorhabditis elegans* cosmid C40D2.  
A:Reference number: Z21323  
A:Accession: T33327  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-147 <MAD>  
A:Cross-references: EMBL:AF077530; PIDN:AAC26267.1; GSPDB:GN00020; CESP:C40D2.4  
A:Experimental source: strain Bristol N2; clone C40D2  
C:Genetics:  
A:Gene: CESP:C40D2.4  
A:Map position: 2

A71518  
probable dnaK suppressor - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C>Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jun-2000  
C:Accession: A71518  
R:Stephens, R.S.; Kaiman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, S.  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
A:Reference number: A71570; MID:99000809  
A:Accession: A71518  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-124 <ARN>  
A:Cross-references: GB:AE001314; GB:AE001273; MID:g332833; PIDN:AAC68004.1; PID:g332883  
C:Genetics:  
A:Gene: dksA  
C:Superfamily: Escherichia coli dosage-dependent dnaK suppressor protein dksA

Query Match 50.0%; Score 5; DB 2; Length 124;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QEVKK 10  
|||||  
Db 31 QEVKK 35

RESULT 20  
dnaK suppressor - Chlamydia pneumoniae (strain CML029)  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jun-2000  
C:Accession: G72066  
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;  
Mature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MID:99206606  
A:Accession: G72066  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-124 <ARN>  
A:Cross-references: GB:AE001638; GB:AE001363; MID:g4376819; PIDN:AAD18674.1; PID:g437682  
A:Experimental source: strain CML029  
C:Genetics:  
A:Gene: dksA  
C:Superfamily: Escherichia coli dosage-dependent dnaK suppressor protein dksA

Query Match 50.0%; Score 5; DB 2; Length 124;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QEVKK 10  
|||||  
Db 31 QEVKK 35

RESULT 21  
hypothetical protein T13F2.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000  
C:Accession: T24876  
R:Swindburne, J.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19947  
A:Accession: T24876  
A>Status: preliminary; translated from GB/EMBL/DBD1  
A:Molecule type: DNA  
A:Residues: 1-124 <MTL>  
A:Cross-references: EMBL:Z81122; PIDN:CAH03353.1; GSPDB:GN00022; CESP:T13F2.2

A:Experimental source: clone T13F2  
C:Genetics:  
A:Gene: CESP:T13F2.2  
A:Map position: 4 61/3; 91/3  
A:introns: 27/3;  
C:Superfamily: Caenorhabditis elegans hypothetical protein T13F2.2

Query Match 50.0%; Score 5; DB 2; Length 124;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QEVKK 10  
|||||  
Db 39 QEVKK 43

RESULT 22  
D86684  
prophage pil protein 41, tail component [imported] - Lactococcus lactis subsp. lactis  
C:Species: Lactococcus lactis subsp. lactis  
C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 23-Mar-2001  
C:Accession: D86684  
R:Boletín, A.; Winkler, P.; Manger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Eh  
Genome Res. in press, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium.  
A:Reference number: A86625  
A:Accession: D86684  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-131 <STO>  
A:Cross-references: GB:AE005176; MID:g12723355; PIDN:AAK04574.1; GSPDB:GN00146  
C:Experimental source: strain IL4403  
C:Genetics:  
A:Gene: pil41

Query Match 50.0%; Score 5; DB 2; Length 131;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PTOEV 8  
|||||  
Db 3 PTOEV 7

RESULT 23  
B81600  
conserved hypothetical protein CP0218 [imported] - Chlamydia pneumoniae (strain A  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Sep-2000  
C:Accession: B81600  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39  
A:Reference number: A81500; MID:20150255  
A:Accession: B81600  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-132 <REA>  
A:Cross-references: GB:AE002183; GB:AE002161; MID:g7189146; PIDN:AAF38087.1; PID:g718  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0218  
C:Superfamily: Escherichia coli dosage-dependent dnaK suppressor protein dksA

Query Match 50.0%; Score 5; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QEVKK 10

A:Reference number: A93891; MUID:82060334  
 A:Accession: A93891  
 A:Molecule type: mRNA  
 A:Residues: 1-103 <HEI>  
 R:Chen, K.C.S.; Kindt, T.J.; Krause, R.M.  
 J. Biol. Chem. 250, 3389-3296, 1975  
 A:Title: Primary structure of the L chain from a rabbit homogeneous antibody to streptococcal antigen A92176; MUID:75133568  
 A:Accession: A92176  
 A:Molecule type: protein  
 A:Residues: 1-57, D, 59-103 <CHE>  
 A:Note: this chain was obtained from antibody to the specific carbohydrate of group C streptococcus. An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chain disulfide bonds. In some cases, such as IGA and IGM, the subunits associate into 16  
 C:Superfamily: Immunoglobulin C region; Immunoglobulin homology  
 C:Keywords: heterotrimer  
 F:19-87/Domain: Immunoglobulin homology <IMM>

Query Match 50.0%; Score 5; DB 1; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
 |||||  
 DB 2 PVAPT 6

RESULT 15  
 F53275  
 Ig kappa-1 chain C region b95 allotype - rabbit (fragment)  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 02-May-1994 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000  
 C:Accession: F53275  
 R:Ayadi, H.; Marche, P.N.; Cazenave, P.A.  
 Immunogenetics 34, 201-207, 1991  
 A:Title: Evolution of the rabbit immunoglobulin kappa chain genes.  
 A:Reference number: A53275; MUID:91372868  
 A:Accession: F53275  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-104 <AYA>  
 A:Note: sequence inconsistent with nucleotide translation  
 A:Note: sequence extracted from NCBI backbone (NCBIN:56069, NCBI:P:56170)  
 C:Superfamily: Immunoglobulin C region; Immunoglobulin homology  
 C:Keywords: heterotrimer; Immunoglobulin  
 F:19-87/Domain: Immunoglobulin homology <IMM>

Query Match 50.0%; Score 5; DB 2; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 82;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
 |||||  
 DB 2 PVAPT 6

RESULT 16  
 T14788  
 hypothetical protein DKFZP564A122.1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T14788  
 R:Duesterhoelt, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wilmann, S.  
 submitted to the Protein Sequence Database, August 1999  
 A:Reference number: 218178  
 A:Accession: T14788  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-112 <DUE>  
 A:Cross-references: EMBL:AL110269  
 A:Experimental source: fetal brain; clone DKFZP564A122

C:Genetics:  
 A:Note: DKFZP564A122.1

Query Match 50.0%; Score 5; DB 2; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 87;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PROEV 8  
 |||||  
 DB 91 PROEV 95

RESULT 17  
 T08411  
 hypothetical protein F18B3.180 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 22-Oct-1999  
 C:Accession: T08411  
 R:Quetier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salan  
 submitted to the Protein Sequence Database, May 1999  
 A:Reference number: 216409  
 A:Accession: T08411  
 A:Molecule type: DNA  
 A:Residues: 1-114 <QUE>  
 A:Cross-references: EMBL:AL049862; GSPDB:GN00061; ATSP:F18B3.180  
 A:Experimental source: cultivar Columbia; BAC clone F18B3  
 C:Genetics:  
 A:Gene: ATSP:F18B3.180  
 A:Map position: 3  
 C:Superfamily: Arabidopsis hypothetical protein F18B3.180

Query Match 50.0%; Score 5; DB 2; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 QEVKK 10  
 |||||  
 DB 43 QEVKK 47

RESULT 18  
 B86557  
 Dnak suppressor [imported] - Chlamydia pneumoniae (strain J138)  
 C:Species: Chlamydia pneumoniae; Chlamydia pneumoniae  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
 C:Accession: B86557  
 R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
 A:Reference number: A86491; MUID:20330349  
 A:Accession: B86557  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-124 <STO>  
 A:Cross-references: GB:BA000008; NID:98978905; PIDN:BA98740.1; GSPDB:GN00142  
 A:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: dksA  
 C:Superfamily: Escherichia coli dosage-dependent dnaK suppressor protein dksA

Query Match 50.0%; Score 5; DB 2; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 QEVKK 10  
 |||||  
 DB 31 QEVKK 35

RESULT 19



Query Match 60.0%; Score 6; DB 2; Length 1133;  
 Best Local Similarity 100.0%; Pred. No. 55;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEYKK 10  
 |||||  
 DB 1099 TOEYKK 1104

## RESULT 10

E64664  
 Outer membrane protein - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999

C:Accession: E64664

R:Tomb, J.F.; White, O.; Karlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalek, H.G.; Glodex, A.; McNamee

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weldman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.

A>Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: E64664

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1230 <TOM>

A:Cross-references: GB:AE000621; GB:AE000511; NID:g2314301; PIDN:AND08205.1; PID:g231431

Query Match 60.0%; Score 6; DB 2; Length 1230;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEYKK 10  
 |||||  
 DB 173 TOEYKK 178

## RESULT 11

E75264

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C:Accession: E75264

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.; Ma

Science 286, 1571-1577, 1999

A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: E75264

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-75 <WHI>

A:Cross-references: GB:AE002081; GB:AE000513; NID:g6460337; PIDN:AAF12064.1; PID:g646034

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2520

A:Map position: 1

Query Match 50.0%; Score 5; DB 2; Length 75;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAPRO 6  
 |||||  
 DB 24 VAPRO 28

## RESULT 12

T29802

hypothetical protein C06E4.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T29802

R:Du, Z.; Gattung, S.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C06E4.

A:Reference number: 220688

A:Accession: T29802

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-82 <DUZ>

A:Cross-references: EMBL:U41277; PIDN:AAA82477.1; CESP:C06E4.2

C:Genetics:

A:Gene: CESP:C06E4.2

A:introns: 31/1

Query Match 50.0%; Score 5; DB 2; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAPRO 6  
 |||||  
 DB 40 VAPRO 44

## RESULT 13

B56581

major basic nuclear protein, variant HCC2 - dihydrofolate (Cryptocodium cohnii) (

N:Alternate names: protein alpha chain p14

C:Species: Cryptocodium cohnii

C>Date: 08-Sep-1995 #sequence\_revision 08-Sep-1995 #text\_change 21-Jul-2000

C:Accession: B56581; S14640; S14644

R:Salas-Rovira, M.; Gerad, M.L.; Caput, D.; Jacques, F.; Soyer-Gobillard, M.O.; Verne

Chromosoma 100, 510-518, 1991

A>Title: Molecular cloning and immunolocalization of two variants of the major basic

A:Reference number: A56581; MUID:92111321

A:Accession: B56581

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-102 <SAL>

A:Cross-references: EMBL:X58444; NID:g17962; PIDN:CAA41350.1; PID:g17963

A>Note: sequence extracted from NCBI backbone (NCBIN:77122, NCBI:77123)

C:Keywords: DNA binding; nucleus

Query Match 50.0%; Score 5; DB 2; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 OEYKK 10  
 |||||  
 DB 43 OEYKK 47

## RESULT 14

KARB

Ig kappa-b4 chain C region - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 31-Jan-1981 #sequence\_revision 15-Oct-1982 #text\_change 16-Aug-1996

C:Accession: A93971; A93891; A92176; A02122

R:Emorine, L.; Dreher, K.; Kindt, T.J.; Max, E.E.

Proc. Natl. Acad. Sci. U.S.A. 80, 5709-5713, 1983

A>Title: Rabbit immunoglobulin kappa genes: structure of a germline b4 allotype J-C 1

A:Reference number: A93971; MUID:83300036

A:Accession: A93971

A:Molecule type: DNA

A:Residues: 1-103 <EMO>

A>Note: the sequence was determined from the germline gene

R:Heidmann, O.; Aufray, C.; Cazenave, P.A.; Rougeon, F.

Proc. Natl. Acad. Sci. U.S.A. 78, 5802-5806, 1981

A>Title: Nucleotide sequence of constant and 3' untranslated regions of a kappa immun

|||||  
Db 29 TOEVKK 34

## RESULT 5

H85574  
hypothetical protein 20891 [imported] - Escherichia coli (strain O157:H7)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
C:Accession: H85574  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dinalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: AB5480; MUID:21074935; PMID:11206551  
A:Accession: H85574  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-456 <STO>  
A:Cross-references: GB:AE005174; NID:q12513654; PIDN:AG55060.1; GSPDB:GN00145; UWGP:Z08  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: 20891

Query Match 60.0%; Score 6; DB 2; Length 456;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VAPTQE 7  
|||||  
Db 378 VAPTQE 383

## RESULT 6

S42719  
actin-binding protein ABP1 - Saccharomyces exiguus  
C:Species: Saccharomyces exiguus  
C:Date: 07-Sep-1994 #sequence\_revision 19-Apr-1996 #text\_change 13-Aug-1999  
C:Accession: S42719  
R:Langre, U.; Stehner, S.; Groll, F.; Wagner, G.; Philippson, P.  
Biochim. Biophys. Acta 1217, 214-218, 1994  
A:Title: Cloning and sequencing of a gene coding for an actin binding protein of Sacchar  
A:Reference number: S42719; MUID:94154001  
A:Accession: S42719  
A:Molecule type: DNA  
A:Residues: 1-617 <LAN>  
A:Cross-references: EMBL:X73977; NID:q433513; PIDN:CAA52156.1; PID:q433514  
C:Superfamily: actin-binding protein ABP1; SH3 homology  
C:Keywords: actin binding; cytoskeleton  
F:564-613/Domain: SH3 homology <SH3>

Query Match 60.0%; Score 6; DB 2; Length 617;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TOEVKK 10  
|||||  
Db 412 TOEVKK 417

## RESULT 7

H84913  
Probable ATP-dependent RNA helicase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: H84913  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Kofel, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487

A:Accession: H84913  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-748 <STO>  
A:Cross-references: GB:AE002093; NID:q2275211; PIDN:AB63833.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g47330  
A:Map position: 2

Query Match 60.0%; Score 6; DB 2; Length 748;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPRQ 6  
|||||  
Db 725 PVAPRQ 730

## RESULT 8

VCLJSP  
env polyprotein - human foamy virus  
N:Alternate names: coat polyprotein  
C:Species: human foamy virus  
A:Note: host Homo sapiens (man)  
C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 16-Jul-1999  
C:Accession: C29685  
R:Fliegel, R.M.; Reithelm, A.; Maurer, B.; Darai, G.  
EMBO J. 6, 2077-2084, 1987  
A:Title: Nucleotide sequence analysis of the env gene and its flanking regions of the  
A:Reference number: A01074; MUID:88004420  
A:Accession: C29685  
A:Molecule type: genomic RNA  
A:Residues: 1-985 <FLU>  
A:Cross-references: GB:X05591; GB:Y00070; NID:q61759; PIDN:CAA29086.1; PID:q61762  
C:Genetics:  
A:Gene: env  
C:Superfamily: foamy virus env polyprotein  
C:Keywords: coat protein; polyprotein; transmembrane protein  
F:64-87/Domain: transmembrane #status predicted <TN1>  
F:579-595/Domain: transmembrane #status predicted <TN2>  
F:936-972/Domain: transmembrane #status predicted <TN3>

Query Match 60.0%; Score 6; DB 1; Length 985;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TOEVKK 10  
|||||  
Db 181 TOEVKK 186

## RESULT 9

T30302  
P-type ATPase - Tetrahymena thermophila  
C:Species: Tetrahymena thermophila  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 07-Dec-1999  
C:Accession: T30302  
R:Wang, S.; Takeyasu, K.  
submitted to the EMBL Data Library, November 1995  
A:Description: The starvation-induced P-type ATPase in Tetrahymena thermophila.  
A:Reference number: Z20815  
A:Accession: T30302  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1133 <MAN>  
A:Cross-references: EMBL:U41063; NID:q1545827; PID:q1545828; PIDN:AB08071.1  
C:Genetics:  
A:Genetic code: SGC5  
A:Note: TPAP

979 4 40.0 163 2 B72225 conserved hypothet  
980 4 40.0 164 2 H71850 flavodoxin - Helic  
981 4 40.0 2 A64655 flavodoxin - Helic  
982 4 40.0 164 2 G02777 durr pyrophosphata  
983 4 40.0 164 2 A26490 histone H1, macron  
984 4 40.0 164 2 S21830 apolipoprotein A-I  
985 4 40.0 164 2 S15641 pillin (variant NGF  
986 4 40.0 164 2 T38684 mitochondrial inne  
987 4 40.0 164 2 B70922 hypothetical prote  
988 4 40.0 164 2 A24659 ZK353.3 protein -  
989 4 40.0 164 2 S18038 homeotic protein S  
990 4 40.0 164 2 H70431 hypothetical prote  
991 4 40.0 165 1 A35702 destrin - chicken  
992 4 40.0 165 2 A27195 histone H1 - Tetra  
993 4 40.0 165 2 S75414 probable ribosomal  
994 4 40.0 165 2 S61230 cytochrome-c biocy  
995 4 40.0 165 2 T39187 probable arylalkyl  
996 4 40.0 165 2 T47011 hypothetical prote  
997 4 40.0 166 1 S12632 cofillin - human  
998 4 40.0 166 1 A29240 cofillin - pig  
999 4 40.0 166 1 S12584 cofillin - mouse  
1000 4 40.0 166 2 B70369 cytochrome c552 -

## ALIGNMENTS

RESULT 1  
A:1487  
protein p60 precursor - Listeria monocytogenes  
N:Alternate names: invasion-associated protein  
C:Species: Listeria monocytogenes  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 13-Sep-1996  
A:Accession: A41487 B41487  
R:Kochner, S.; Leimaster-Waechter, M.; Chakraborty, T.; Lottspeich, F.; Goebel, W.  
Infect. Immun. 58, 1943-1950, 1990  
A:Title: The gene coding for protein p60 of Listeria monocytogenes and its use as a spec  
A:Reference number: A41487; MUID:90256283  
A:Accession: A41487  
A:Molecule type: DNA  
A:Residues: 1484 <KOE>  
A:Cross-references: GB:X52268  
A:Accession: B41487  
A:Molecule type: protein  
A:Residues: 28-49 <KO2>  
C:Genetics:  
A:Gene: lap  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-484/Product: protein 60 #status predicted <MAT>

Query Match 100.0%; Score 10; DB 2; Length 484;  
Best Local Similarity 100.0%; Pred. No. 0.0013;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPTOEVKK 10  
|||||  
Db 148 PVAPTOEVKK 157

RESULT 2  
D81259  
probable long transport protein Cj1630 (imported) - Campylobacter jejuni (strain NCRC 11  
C:Species: Campylobacter jejuni  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 31-Mar-2000  
C:Accession: D81259  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barre  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals h  
A:Reference number: A81250; MUID:20150912  
A:Accession: D81259  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-227 <PAR>  
A:Cross-references: GB:AL139079; GB:AL111168; NID:96968971; PIDN:CAB72618.1; PID:9696  
A:Experimental source: serotype O2, strain NCRC 11168  
C:Genetics:  
A:Gene: tonB2; Cj1630

Query Match 70.0%; Score 7; DB 2; Length 227;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APTOEVK 9  
|||||  
Db 105 APTOEVK 111

RESULT 3  
F83231  
hypothetical protein PA3307 (imported) - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83231  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A:Reference number: A82950; MUID:20437337  
A:Accession: F83231  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-102 <SNO>  
A:Cross-references: GB:AE004753; GB:AE004091; NID:99949433; PIDN:AG06695.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3307

Query Match 60.0%; Score 6; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 6.6;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPTQ 6  
|||||  
Db 27 PVAPTQ 32

RESULT 4  
T37186  
probable regulatory protein - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T37186  
R:Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
Submitted to the EMBL Data Library, July 1999  
A:Reference number: Z21598  
A:Accession: T37186  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-116 <SEE>  
A:Cross-references: EMBL:AL096823; PIDN:CAB46971.1; GSPDB:GN00070; SCOEDB:SCQ11.16  
C:Genetics:  
A:Gene: SCOEDB:SCQ11.16  
C:Superfamily: Streptomyces ambifaciens replication activator protein pra

Query Match 60.0%; Score 6; DB 2; Length 116;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TOEVKK 10

833	4	40.0	142	2	E71079	hypothetical prote	906	4	40.0	153	2	S59125	histone H2B [valid
834	4	40.0	142	2	T51479	hypothetical prote	907	4	40.0	153	2	T07624	heat shock protein
835	4	40.0	143	2	G84471	En/Spm-like transp	908	4	40.0	153	2	B42748	trans-regulatory s
836	4	40.0	143	2	T03231	hypothetical prote	909	4	40.0	153	2	S29277	hypothetical prote
837	4	40.0	143	2	T08789	hypothetical prote	910	4	40.0	153	2	E71229	probable purine ph
838	4	40.0	144	1	A61632	granulocyte-macrop	911	4	40.0	153	2	A75197	purine phosphoribo
839	4	40.0	144	1	SMXD12	hexon-associated p	912	4	40.0	153	2	JC5854	polyketide synthas
840	4	40.0	144	2	JH0469	granulocyte-macrop	913	4	40.0	154	2	S12629	heat shock cognate
841	4	40.0	144	2	PH1563	Ig heavy chain v r	914	4	40.0	154	2	T07625	heat shock protein
842	4	40.0	144	2	E41287	Ig heavy chain pre	915	4	40.0	154	2	T07629	small heat shock p
843	4	40.0	144	2	A36335	ribosomal protein	916	4	40.0	154	2	B6105	hypothetical prote
844	4	40.0	144	2	E75128	glycine cleavage s	917	4	40.0	154	2	A72029	conserved hypoteth
845	4	40.0	144	2	S66570	biotin carboxyl ca	918	4	40.0	154	2	H86594	CT556 hypothetical
846	4	40.0	144	2	B70404	transcription regu	919	4	40.0	154	2	A85551	hypothetical prote
847	4	40.0	144	2	H70971	hypothetical prote	920	4	40.0	154	2	T17580	hypothetical prote
848	4	40.0	144	2	C75376	hypothetical prote	921	4	40.0	154	2	C82207	hypothetical prote
849	4	40.0	144	2	UC2102	lectin-related pro	922	4	40.0	154	2	PC4031	hypothetical 154 p
850	4	40.0	144	2	T46334	hypothetical prote	923	4	40.0	154	2	S55017	hypothetical prote
851	4	40.0	144	2	B75284	Mutr/ndix family	924	4	40.0	154	2	F72741	hypothetical prote
852	4	40.0	145	2	A83843	peptidyl-prolyl ci	925	4	40.0	155	1	SNRB2	pathogenesis-relat
853	4	40.0	145	2	H69082	ribosomal protein	926	4	40.0	155	2	S59956	DNA-directed RNA p
854	4	40.0	145	2	F69214	heat shock protein	927	4	40.0	155	2	S21659	ribonuclease H (EC
855	4	40.0	145	2	S04042	embryonic abundant	928	4	40.0	155	2	T06449	probable heat shoc
856	4	40.0	145	2	C70328	conserved hypoteth	929	4	40.0	155	2	T11670	pathogenesis relat
857	4	40.0	145	2	S63134	probable membrane	930	4	40.0	155	2	E64105	acetyl-CoA carboxy
858	4	40.0	146	2	S60700	gag protein - huma	931	4	40.0	155	2	H75302	conserved hypoteth
859	4	40.0	146	2	S60697	gag protein - huma	932	4	40.0	155	2	A82884	conserved hypoteth
860	4	40.0	146	2	S60699	gag protein - huma	933	4	40.0	156	1	SNRBI	pathogenesis-relat
861	4	40.0	146	2	S60708	gag protein - huma	934	4	40.0	156	1	C64390	probable transcrip
862	4	40.0	146	2	G75544	hypothetical prote	935	4	40.0	156	2	S71566	heat shock protein
863	4	40.0	146	2	T49867	hypothetical prote	936	4	40.0	156	2	D81343	probable integral
864	4	40.0	146	2	B71430	hypothetical prote	937	4	40.0	157	2	D70541	hypothetical prote
865	4	40.0	147	2	PH1561	Ig heavy chain v r	938	4	40.0	157	2	S04939	heat shock 22k pro
866	4	40.0	147	2	A58345	myoglobin - tremet	939	4	40.0	157	2	S31634	hypothetical prote
867	4	40.0	147	2	C47056	cnr regulatory pro	940	4	40.0	157	2	F71406	hypothetical prote
868	4	40.0	147	2	T49624	hypothetical prote	941	4	40.0	158	2	S16247	heat shock protein
869	4	40.0	147	2	D32804	chorismate mutase	942	4	40.0	158	2	S20518	hypothetical prote
870	4	40.0	147	2	S70833	methylated-DNA--pr	943	4	40.0	158	2	S20517	hypothetical prote
871	4	40.0	147	2	A42366	flagellar basal bo	944	4	40.0	158	2	B75141	hypothetical prote
872	4	40.0	147	2	T02223	hypothetical prote	945	4	40.0	158	2	D75069	hypothetical prote
873	4	40.0	147	2	T26225	hypothetical prote	946	4	40.0	159	2	S72544	heat shock protein
874	4	40.0	147	2	B70436	hypothetical prote	947	4	40.0	159	2	C72210	conserved hypoteth
875	4	40.0	148	2	S29257	Ig heavy chain v r	948	4	40.0	159	2	C64547	hypothetical prote
876	4	40.0	148	2	G83613	probable acetylra	949	4	40.0	159	2	T48837	hypothetical prote
877	4	40.0	148	2	A70077	conserved hypoteth	950	4	40.0	159	2	D70133	conserved hypoteth
878	4	40.0	148	2	T75205	molybdopterin bios	951	4	40.0	159	2	T17768	hypothetical prote
879	4	40.0	148	2	T05143	hypothetical prote	952	4	40.0	159	2	T46954	hypothetical prote
880	4	40.0	148	2	T21488	hypothetical prote	953	4	40.0	160	2	S78308	peptidylprolyl iso
881	4	40.0	148	2	A86878	non-heme iron-bind	954	4	40.0	160	2	T05740	heat shock protein
882	4	40.0	149	2	D71929	ribosomal protein	955	4	40.0	160	2	H69015	ribosomal protein
883	4	40.0	149	2	E69653	transcription regu	956	4	40.0	160	2	D75461	hypothetical prote
884	4	40.0	149	2	F71240	molybdopterin bios	957	4	40.0	160	2	F84350	hypothetical prote
885	4	40.0	149	2	S49527	zona pellucida bin	958	4	40.0	161	2	E81211	hypothetical prote
886	4	40.0	150	1	MORTA2	myosin alkali ligh	959	4	40.0	161	2	S44757	heat shock protein
887	4	40.0	150	1	MORBL2	myosin alkali ligh	960	4	40.0	161	2	S00646	heat shock protein
888	4	40.0	150	1	T25581	hypothetical prote	961	4	40.0	161	2	S43260	helix-loop-helix p
889	4	40.0	150	2	B64584	ribosomal protein	962	4	40.0	161	2	D82042	transcription elon
890	4	40.0	150	2	S20874	heat shock protein	963	4	40.0	161	2	C36810	hypothetical prote
891	4	40.0	150	2	B23253	myosin A2 catalyti	964	4	40.0	161	2	H82782	hypothetical prote
892	4	40.0	150	2	G72660	hypothetical prote	965	4	40.0	161	2	E69328	3-isopropylmalate
893	4	40.0	151	1	G84173	archaeal histone A	966	4	40.0	161	2	T23912	hypothetical prote
894	4	40.0	151	1	HHWT17	heat shock protein	967	4	40.0	162	2	B72371	hypothetical prote
895	4	40.0	151	2	S21600	heat shock protein	968	4	40.0	162	2	G96671	hypothetical prote
896	4	40.0	151	2	T05739	probable heat shoc	969	4	40.0	163	1	YCEC3H	acetoacetate synth
897	4	40.0	151	2	T42298	hypothetical prote	970	4	40.0	163	2	S15940	acetoacetate synth
898	4	40.0	151	2	F82427	hypothetical prote	971	4	40.0	163	2	B64131	acetoacetate synth
899	4	40.0	152	2	S08350	superoxide dismuta	972	4	40.0	163	2	B85490	hypothetical prote
900	4	40.0	152	2	B71119	probable aspartate	973	4	40.0	163	2	F83059	acetoacetate synth
901	4	40.0	152	2	C72227	conserved hypoteth	974	4	40.0	163	2	H81066	acetoacetate synth
902	4	40.0	152	2	T49126	hypothetical prote	975	4	40.0	163	2	S71567	small heat-shock p
903	4	40.0	153	1	MYTWM	myoglobin - map tu	976	4	40.0	163	2	A72762	probable bacteriof
904	4	40.0	153	1	MYTNG	myoglobin - green	977	4	40.0	163	2	D71062	hypothetical prote
905	4	40.0	153	1	WRBPL5	early protein gpl7	978	4	40.0	163	2	B96842	hypothetical prote

687	4	40.0	124	1	TIBMB	trypsin inhibitor	760	4	40.0	132	2	S31596	Ig heavy chain V r
688	4	40.0	124	1	M1HWM	Ig heavy chain V-I	761	4	40.0	132	2	S46394	Ig heavy chain V r
689	4	40.0	124	1	M1HUS1	Ig heavy chain V-I	762	4	40.0	132	2	S49160	hypothetical prote
690	4	40.0	124	1	S19665	Ig heavy chain V r	763	4	40.0	132	2	B70312	hypothetical prote
691	4	40.0	124	2	F72403	glycine cleavage s	764	4	40.0	132	2	S30969	gene 24 protein -
692	4	40.0	124	2	B70472	flagellar protein	765	4	40.0	132	2	T139773	hypothetical prote
693	4	40.0	124	2	G75116	hypothetical prote	766	4	40.0	133	1	SXADP1	hexon-associated p
694	4	40.0	124	2	E82196	RstB1 protein VC14	767	4	40.0	133	2	S73769	adenine phosphorib
695	4	40.0	124	2	T36292	hypothetical prote	768	4	40.0	133	2	C33548	Ig heavy chain V-1
696	4	40.0	125	1	S1PSPT	steroid Delta-1som	769	4	40.0	133	2	T31427	profilin - common
697	4	40.0	125	1	HVHUMO	Ig heavy chain V-I	770	4	40.0	133	2	A60758	17k antigen - nema
698	4	40.0	125	1	PH1410	Ig heavy chain V r	771	4	40.0	134	2	S54828	cysteine proteinas
699	4	40.0	125	2	PH0957	Ig heavy chain V r	772	4	40.0	134	2	JC4882	cystatin - maize
700	4	40.0	125	2	A64502	hypothetical prote	773	4	40.0	134	2	PH1422	Ig heavy chain V r
701	4	40.0	125	2	S78697	probable export pr	774	4	40.0	134	2	T22275	hypothetical prote
702	4	40.0	125	2	H86722	hypothetical prote	775	4	40.0	134	2	F72802	gp24 protein - Myc
703	4	40.0	125	2	T38537	probable single-st	776	4	40.0	134	2	F75416	hypothetical prote
704	4	40.0	126	2	PH1417	Ig heavy chain V r	777	4	40.0	135	2	JC4007	cystatin II - maiz
705	4	40.0	126	2	PH1418	Ig heavy chain V r	778	4	40.0	135	2	S27239	cysteine proteinas
706	4	40.0	126	2	PH1424	Ig heavy chain V r	779	4	40.0	135	2	PH0953	Ig heavy chain V r
707	4	40.0	126	2	PH1412	Ig heavy chain V r	780	4	40.0	135	2	B32274	Ig heavy chain pre
708	4	40.0	126	2	B33548	Ig heavy chain V-1	781	4	40.0	135	2	S49530	anti-Sm antibody V
709	4	40.0	126	2	A56657	PEMW2/MESA protei	782	4	40.0	135	2	T09876	denhyrin - upland
710	4	40.0	126	2	G83571	conserved hypochet	783	4	40.0	135	2	C64453	translatation initia
711	4	40.0	126	2	H75433	hypothetical prote	784	4	40.0	135	2	A70317	hypothetical prote
712	4	40.0	127	1	C64140	probable glycy l ra	785	4	40.0	135	2	T40119	transfer complex p
713	4	40.0	127	2	PH1414	Ig heavy chain V r	786	4	40.0	135	2	D36891	hypothetical prote
714	4	40.0	127	2	PH1415	Ig heavy chain V r	787	4	40.0	135	2	G722714	hypothetical prote
715	4	40.0	127	2	PH1421	Ig heavy chain V r	788	4	40.0	136	2	PH1559	Ig heavy chain V r
716	4	40.0	127	2	PH1411	Ig heavy chain V r	789	4	40.0	136	2	A49047	Ig heavy chain V r
717	4	40.0	127	2	PH1420	Ig heavy chain V r	790	4	40.0	136	2	S31600	Ig heavy chain V r
718	4	40.0	127	2	PH0955	Ig heavy chain V r	791	4	40.0	136	2	PH0960	Ig heavy chain V r
719	4	40.0	127	2	S34014	Ig heavy chain V r	792	4	40.0	136	2	S46102	hypothetical prote
720	4	40.0	127	2	R82197	RstB2 protein VC14	793	4	40.0	136	2	T133388	hypothetical prote
721	4	40.0	128	2	T69024	MHC sex-limited pr	794	4	40.0	136	2	B81388	hypothetical prote
722	4	40.0	128	2	S16685	Ig heavy chain V r	795	4	40.0	136	2	E70182	hypothetical prote
723	4	40.0	128	2	PH0952	Ig heavy chain V r	796	4	40.0	137	2	PH1562	Ig heavy chain V r
724	4	40.0	128	2	S76468	hypothetical prote	797	4	40.0	137	2	T22043	hypothetical prote
725	4	40.0	128	2	T22276	hypothetical prote	798	4	40.0	137	2	H70472	conserved hypochet
726	4	40.0	128	2	D75575	5-carboxymethyl-2-	799	4	40.0	137	2	S74770	hypothetical prote
727	4	40.0	128	2	S76955	hypothetical prote	800	4	40.0	137	2	A05005	ribulose-bisphosph
728	4	40.0	128	2	D72481	hypothetical prote	801	4	40.0	138	2	PN0538	Ig heavy chain V r
729	4	40.0	128	2	H70435	hypothetical prote	802	4	40.0	138	2	PH1564	Ig heavy chain V r
730	4	40.0	128	2	T15911	hypothetical prote	803	4	40.0	138	2	PH1565	Ig heavy chain V r
731	4	40.0	128	2	S75366	hypothetical prote	804	4	40.0	138	2	S24078	envelope protein -
732	4	40.0	128	2	T32789	hypothetical prote	805	4	40.0	138	2	D83833	hypothetical prote
733	4	40.0	129	2	A33548	Ig heavy chain V-1	806	4	40.0	138	2	B45244	transcription regu
734	4	40.0	129	2	S36260	Ig heavy chain V r	807	4	40.0	138	2	T04560	hypothetical prote
735	4	40.0	129	2	S46393	Ig heavy chain V r	808	4	40.0	138	2	B83366	hypothetical prote
736	4	40.0	129	2	B71350	probable ribosomal	809	4	40.0	138	2	G85806	hypothetical prote
737	4	40.0	129	2	S53806	ribosomal protein	810	4	40.0	138	2	T29558	hypothetical prote
738	4	40.0	129	2	S37713	hypothetical prote	811	4	40.0	138	2	T47152	hypothetical prote
739	4	40.0	129	2	F83764	ectoine synthase	812	4	40.0	139	2	PH1558	Ig heavy chain V r
740	4	40.0	130	1	F64376	hypothetical prote	813	4	40.0	139	2	C70114	ribosomal protein
741	4	40.0	130	1	C69883	ribonucleoprotein	814	4	40.0	139	2	T34633	hypothetical prote
742	4	40.0	130	1	PC6025	cystine proteinas	815	4	40.0	139	2	G69497	iron-dependent rep
743	4	40.0	130	2	S08079	Ig kappa chain pre	816	4	40.0	139	2	T23471	hypothetical prote
744	4	40.0	130	2	F25733	T-cell receptor al	817	4	40.0	140	2	H69526	methylmalonyl-CoA
745	4	40.0	130	2	G70813	probable lipqS proc	818	4	40.0	140	2	H64664	conserved hypochet
746	4	40.0	130	2	S77271	hypothetical prote	819	4	40.0	140	2	B69323	hypothetical prote
747	4	40.0	130	2	D70966	probable phage pro	820	4	40.0	140	2	C69269	hypothetical prote
748	4	40.0	130	2	F64348	hypothetical prote	821	4	40.0	141	2	A46256	dUTP pyrophosphata
749	4	40.0	130	2	D84363	hypothetical prote	822	4	40.0	141	2	S31685	Ig heavy chain V r
750	4	40.0	130	2	T51879	hypothetical prote	823	4	40.0	141	2	G71099	hypothetical prote
751	4	40.0	131	1	CCTW5T	cytochrome c552 [v	824	4	40.0	141	2	S34629	hypothetical prote
752	4	40.0	131	2	PH1425	Ig heavy chain V r	825	4	40.0	141	2	H84059	hypothetical prote
753	4	40.0	131	2	S21924	Ig heavy chain V r	826	4	40.0	141	2	T48839	hypothetical prote
754	4	40.0	131	2	T35341	hypothetical prote	827	4	40.0	141	2	T15990	hypothetical prote
755	4	40.0	131	2	G75208	hypothetical prote	828	4	40.0	142	1	S42722	interleukin-3 prec
756	4	40.0	131	2	G86802	prophage p13 prote	829	4	40.0	142	1	S42721	interleukin-3 prec
757	4	40.0	132	1	E70434	conserved hypochet	830	4	40.0	142	2	E70356	hypothetical prote
758	4	40.0	132	1	PH1427	Ig heavy chain V r	831	4	40.0	142	2	A32483	Ig heavy chain V r
759	4	40.0	132	2	PH0954	Ig heavy chain V r	832	4	40.0	142	2	D69510	oxaloacetate decar

541	4	40.0	107	2	S40290	protein-tyrosine-p	614	4	40.0	116	2	PN0618	ribosomal protein
542	4	40.0	107	2	S76718	hypothetical prote	615	4	40.0	116	2	JC5188	sperulation-specif
543	4	40.0	107	2	T45621	hypothetical prote	616	4	40.0	116	2	PC4228	N-acetylglucosamin
544	4	40.0	107	2	E83182	hypothetical prote	617	4	40.0	116	2	T33195	hypothetical prote
545	4	40.0	107	2	E83040	DNA-binding protei	618	4	40.0	116	2	H72489	hypothetical prote
546	4	40.0	107	2	G81080	hypothetical prote	619	4	40.0	117	1	G1H0EU	Ig heavy chain V-I
547	4	40.0	108	2	PH1664	Ig heavy chain V r	620	4	40.0	117	1	HVH0UG	Ig heavy chain pre
548	4	40.0	108	2	C49056	T-cell receptor al	621	4	40.0	117	1	HVH0UG	Ig heavy chain pre
549	4	40.0	108	2	T51146	ring-box protein 1	622	4	40.0	117	2	T02872	probable lipid tra
550	4	40.0	108	2	A86675	hypothetical prote	623	4	40.0	117	2	S19669	Ig heavy chain V r
551	4	40.0	108	2	S61627	hypothetical prote	624	4	40.0	117	2	S19670	Ig heavy chain V r
552	4	40.0	108	2	D72662	hypothetical prote	625	4	40.0	117	2	A28846	Ig heavy chain pre
553	4	40.0	108	2	A72516	hypothetical prote	626	4	40.0	117	2	PT0371	Ig gamma chain pre
554	4	40.0	108	2	T26681	hypothetical prote	627	4	40.0	117	2	S21668	Ig kappa chain V r
555	4	40.0	108	2	A72704	hypothetical prote	628	4	40.0	117	2	S31680	Ig heavy chain V r
556	4	40.0	109	2	S16437	ribulose-bisphosph	629	4	40.0	117	2	S18551	Ig heavy chain V r
557	4	40.0	109	2	PH1672	Ig heavy chain V r	630	4	40.0	117	2	S18552	Ig heavy chain V r
558	4	40.0	109	2	PH1668	Ig heavy chain V r	631	4	40.0	117	2	S18553	Ig heavy chain V r
559	4	40.0	109	2	PH1671	Ig heavy chain V r	632	4	40.0	117	2	B82614	hypothetical prote
560	4	40.0	109	2	D71279	hypothetical prote	633	4	40.0	118	2	H66640	ribonuclease P (EC
561	4	40.0	109	2	F84010	hypothetical prote	634	4	40.0	118	2	PH1666	Ig heavy chain V r
562	4	40.0	109	2	T51864	probable heat-shoc	635	4	40.0	118	2	S36265	Ig heavy chain V r
563	4	40.0	110	1	A37414	parvalbumin - mous	636	4	40.0	118	2	S12440	Ig heavy chain V r
564	4	40.0	110	2	PH1669	Ig heavy chain V r	637	4	40.0	118	2	S12441	Ig lambda chain (M
565	4	40.0	110	2	PH1670	Ig heavy chain V r	638	4	40.0	118	2	S12442	Ig lambda chain (K
566	4	40.0	110	2	G70305	Ig heavy chain V r	639	4	40.0	118	2	S12442	Ig lambda chain (K
567	4	40.0	110	2	S50100	ribosomal protein	640	4	40.0	118	2	T07905	hypothetical prote
568	4	40.0	110	2	A55263	transcription-asso	641	4	40.0	119	2	S08077	low-carbon dioxide
569	4	40.0	110	2	JV0102	hypothetical 12.8k	642	4	40.0	119	2	PN0295	Ig kappa chain pre
570	4	40.0	110	2	T30125	hypothetical prote	643	4	40.0	119	2	PH0961	Ig heavy chain V-D
571	4	40.0	110	2	D70610	probable pg protei	644	4	40.0	119	2	S44106	Ig heavy chain V r
572	4	40.0	110	2	T27823	hypothetical prote	645	4	40.0	119	2	S58445	Ig heavy chain V-D
573	4	40.0	110	2	T46071	hypothetical prote	646	4	40.0	119	2	T00151	lipoprotein D - Sa
574	4	40.0	110	2	T48382	hypothetical prote	647	4	40.0	119	2	E83674	hypothetical prote
575	4	40.0	111	2	S21925	Ig heavy chain V r	648	4	40.0	119	2	G86843	hypothetical prote
576	4	40.0	111	2	S17200	protein kinase (EC	649	4	40.0	119	2	D64002	hypothetical prote
577	4	40.0	111	2	D72646	hypothetical prote	650	4	40.0	120	2	S21667	Ig kappa chain V r
578	4	40.0	112	2	S77344	carbon dioxide con	651	4	40.0	120	2	S21666	Ig kappa chain V r
579	4	40.0	112	2	A71515	hypothetical prote	652	4	40.0	120	2	PH0962	Ig heavy chain V r
580	4	40.0	112	2	A64976	galactitol utilisaz	653	4	40.0	120	2	T29774	hypothetical prote
581	4	40.0	112	2	C83938	hypothetical prote	654	4	40.0	120	2	D71039	hypothetical prote
582	4	40.0	112	2	C86415	hypothetical prote	655	4	40.0	120	2	C85849	unknown protein en
583	4	40.0	113	1	RKWT55	ribulose-bisphosph	656	4	40.0	120	2	A72539	hypothetical prote
584	4	40.0	113	1	R6D0P1	acidic ribosomal p	657	4	40.0	121	2	E70313	histidine triad-11
585	4	40.0	113	2	PH1428	Ig heavy chain V r	658	4	40.0	121	2	A28678	T-cell receptor ga
586	4	40.0	113	2	PH1663	Ig heavy chain V r	659	4	40.0	121	2	S20783	Ig heavy chain V r
587	4	40.0	113	2	PC1281	NS5 protein - hepa	660	4	40.0	121	2	A49590	Ig heavy chain V r
588	4	40.0	113	2	S11640	hypothetical prote	661	4	40.0	121	2	D75089	hypothetical prote
589	4	40.0	113	2	A56581	major basic nuclea	662	4	40.0	121	2	T32888	hypothetical prote
590	4	40.0	113	2	A83429	hypothetical prote	663	4	40.0	121	2	S74021	hypothetical prote
591	4	40.0	113	2	D83220	hypothetical prote	664	4	40.0	121	2	S75660	hypothetical prote
592	4	40.0	113	2	A83697	hypothetical prote	665	4	40.0	122	2	PC4279	ant1-SS-A/Ro 60k p
593	4	40.0	113	2	T08623	probable sensor ki	666	4	40.0	122	2	PC4280	ant1-SS-A/Ro 60k p
594	4	40.0	113	2	S66805	hypothetical prote	667	4	40.0	122	2	PH1426	Ig heavy chain V r
595	4	40.0	113	2	T49533	hypothetical prote	668	4	40.0	122	2	PH0958	Ig heavy chain V r
596	4	40.0	113	2	PC2231	hobo-like transpos	669	4	40.0	122	2	S36271	Ig heavy chain V r
597	4	40.0	114	2	S07898	endothelin 3 - tab	670	4	40.0	122	2	C49590	Ig heavy chain V r
598	4	40.0	114	2	PH1667	Ig heavy chain V r	671	4	40.0	122	2	B49590	Ig heavy chain V r
599	4	40.0	115	1	JN0318	guanylin precursor	672	4	40.0	122	2	T22096	Ig heavy chain V r
600	4	40.0	115	2	PH1560	Ig heavy chain V r	673	4	40.0	123	1	RKSPS	hypothetical prote
601	4	40.0	115	2	PH1557	Ig heavy chain V r	674	4	40.0	123	2	D33548	ribulose-bisphosph
602	4	40.0	115	2	G72566	hypothetical prote	675	4	40.0	123	2	S44108	Ig heavy chain V-1
603	4	40.0	115	2	E70702	hypothetical prote	676	4	40.0	123	2	PH1413	Ig heavy chain V r
604	4	40.0	115	2	E86809	hypothetical prote	677	4	40.0	123	2	PH1423	Ig heavy chain V r
605	4	40.0	115	2	E84501	hypothetical prote	678	4	40.0	123	2	S38492	Ig heavy chain - h
606	4	40.0	115	2	E84512	probable MYB fam11	679	4	40.0	123	2	C36006	Ig heavy chain V r
607	4	40.0	115	2	PC2230	hobo-like transpos	680	4	40.0	123	2	E64418	Ig heavy chain V-D
608	4	40.0	115	2	BA9094	methylmalonyl-CoA	681	4	40.0	123	2	C71143	conserved hypotet
609	4	40.0	116	1	B46279	guanylin precursor	682	4	40.0	123	2	H75059	hypothetical prote
610	4	40.0	116	2	S31698	Ig heavy chain pre	683	4	40.0	123	2	B75546	hypothetical prote
611	4	40.0	116	2	PH0959	Ig heavy chain V r	684	4	40.0	123	2	C71423	hypothetical prote
612	4	40.0	116	2	S36261	Ig heavy chain V r	685	4	40.0	123	2	A05125	hypothetical prote
613	4	40.0	116	2	S31667	Ig heavy chain V r	686	4	40.0	123	2	T20279	hypothetical prote

395	4	40.0	81	2	T29888	hypothetical prote
396	4	40.0	82	2	S37173	ribosomal protein
397	4	40.0	82	2	G82365	tata protein WC008
398	4	40.0	82	2	T29889	hypothetical prote
399	4	40.0	82	2	T29893	hypothetical prote
400	4	40.0	82	2	T29891	hypothetical prote
401	4	40.0	82	2	T33088	hypothetical prote
402	4	40.0	83	2	C47188	MHC class II histo
403	4	40.0	83	2	T08238	gas-vesicle operon
404	4	40.0	83	2	J01124	gas-vesicle operon
405	4	40.0	84	2	A33012	serum response fac
406	4	40.0	84	2	S61465	p83/100 protein -
407	4	40.0	84	2	S61466	p83/100 protein -
408	4	40.0	84	2	T47420	hypothetical prote
409	4	40.0	84	2	E82709	hypothetical prote
410	4	40.0	86	2	F85748	unknown protein en
411	4	40.0	86	2	A70005	conserved hypothet
412	4	40.0	87	2	T27141	hypothetical prote
413	4	40.0	87	2	S27040	VAT-1 protein - Pa
414	4	40.0	88	1	O6BP77	gene 6.7 protein -
415	4	40.0	88	2	S60815	M protein precursor
416	4	40.0	88	2	A57399	EM12 protein - Ba
417	4	40.0	88	2	A83659	hypothetical prote
418	4	40.0	89	2	B69352	acylphosphatase (a
419	4	40.0	90	2	S78047	DNA-directed RNA p
420	4	40.0	90	2	B71983	hypothetical prote
421	4	40.0	91	2	B96574	hypothetical prote
422	4	40.0	92	2	C82174	ribosomal protein
423	4	40.0	92	2	F84620	similar to late em
424	4	40.0	92	2	T34744	hypothetical prote
425	4	40.0	92	2	T51187	small zinc finger-
426	4	40.0	92	2	JU0239	nuclear matrix pro
427	4	40.0	93	2	B85537	unknown [imported]
428	4	40.0	93	2	T45593	small zinc finger-
429	4	40.0	94	2	S49470	hypothetical prote
430	4	40.0	94	2	B84539	hypothetical prote
431	4	40.0	94	2	G36943	chemoreceptor prot
432	4	40.0	94	2	S43903	hypothetical prote
433	4	40.0	95	2	S68231	PHG22 protein prec
434	4	40.0	95	2	B32830	homeotic protein H
435	4	40.0	95	2	A56644	inverted repeat co
436	4	40.0	95	2	H81119	conserved hypothet
437	4	40.0	95	2	G81901	hypothetical prote
438	4	40.0	95	2	C69905	hypothetical prote
439	4	40.0	96	2	S10069	regulatory protein
440	4	40.0	96	1	USPSBM	small acid-soluble
441	4	40.0	97	2	PH0870	Ig heavy chain V r
442	4	40.0	97	2	I51262	eggreacan chondroit
443	4	40.0	97	2	A69061	Ig heavy chain V r
444	4	40.0	98	2	PH1429	Ig heavy chain V r
445	4	40.0	98	2	PH1274	Ig heavy chain V r
446	4	40.0	98	2	PH0877	Ig heavy chain V r
447	4	40.0	98	2	A30523	Ig heavy chain V r
448	4	40.0	98	2	PH0871	Ig heavy chain V-1
449	4	40.0	98	2	S24680	Ig heavy chain V r
450	4	40.0	98	2	S46463	Ig heavy chain VI
451	4	40.0	98	2	S26910	Ig heavy chain V r
452	4	40.0	98	2	S26914	Ig heavy chain V r
453	4	40.0	98	2	S26918	Ig heavy chain V r
454	4	40.0	98	2	S26938	Ig heavy chain V r
455	4	40.0	98	2	S26909	Ig heavy chain V r
456	4	40.0	98	2	S26919	Ig heavy chain V r
457	4	40.0	98	2	S26913	Ig heavy chain V r
458	4	40.0	98	2	S26912	Ig heavy chain V r
459	4	40.0	98	2	S26911	Ig heavy chain V r
460	4	40.0	98	2	S26920	Ig heavy chain V r
461	4	40.0	98	2	S26921	Ig heavy chain V r
462	4	40.0	98	2	S26915	Ig heavy chain V r
463	4	40.0	98	2	S26907	Ig heavy chain V r
464	4	40.0	98	2	S22592	hypothetical prote
465	4	40.0	98	2	T24286	hypothetical prote
466	4	40.0	99	1	R6MKL2	ribosomal protein
467	4	40.0	99	1	Q00YBW	hypothetical prote
468	4	40.0	99	2	H69000	translation initia
469	4	40.0	100	2	T08763	hypothetical prote
470	4	40.0	100	2	B49598	hypothetical prote
471	4	40.0	100	2	S77878	hypothetical prote
472	4	40.0	101	2	S12424	Ig heavy chain V r
473	4	40.0	101	2	S12428	Ig heavy chain V r
474	4	40.0	101	2	S12431	Ig heavy chain V r
475	4	40.0	101	2	H65009	hypothetical prote
476	4	40.0	101	2	B72079	hypothetical prote
477	4	40.0	101	2	F86544	hypothetical prote
478	4	40.0	101	2	S76550	hypothetical prote
479	4	40.0	102	1	W7WLEP	E7 protein - Europ
480	4	40.0	102	2	PH1232	Ig heavy chain V r
481	4	40.0	102	2	PH1234	Ig heavy chain V r
482	4	40.0	102	2	PH1235	Ig heavy chain V r
483	4	40.0	102	2	PH1237	Ig heavy chain V r
484	4	40.0	102	2	PH1238	Ig heavy chain V r
485	4	40.0	102	2	PH1239	Ig heavy chain V r
486	4	40.0	102	2	PH1240	Ig heavy chain V r
487	4	40.0	102	2	PH1243	Ig heavy chain V r
488	4	40.0	102	2	PH1248	Ig heavy chain V r
489	4	40.0	102	2	PH1262	Ig heavy chain V r
490	4	40.0	102	2	PH1264	Ig heavy chain V r
491	4	40.0	102	2	PH1268	Ig heavy chain V r
492	4	40.0	102	2	PH1276	Ig heavy chain V r
493	4	40.0	102	2	PH1280	Ig heavy chain V r
494	4	40.0	102	2	PH1241	Ig heavy chain V r
495	4	40.0	102	2	PH1244	Ig heavy chain V r
496	4	40.0	102	2	PH1247	Ig heavy chain V r
497	4	40.0	102	2	PH1249	Ig heavy chain V r
498	4	40.0	102	2	PH1252	Ig heavy chain V r
499	4	40.0	102	2	PH1254	Ig heavy chain V r
500	4	40.0	102	2	PH1258	Ig heavy chain V r
501	4	40.0	102	2	PH1259	Ig heavy chain V r
502	4	40.0	102	2	PH1260	Ig heavy chain V r
503	4	40.0	102	2	PH1263	Ig heavy chain V r
504	4	40.0	102	2	PH1265	Ig heavy chain V r
505	4	40.0	102	2	PH1266	Ig heavy chain V r
506	4	40.0	102	2	PH1267	Ig heavy chain V r
507	4	40.0	102	2	PH1275	Ig heavy chain V r
508	4	40.0	102	2	PH1278	Ig heavy chain V r
509	4	40.0	102	2	PH1279	Ig heavy chain V r
510	4	40.0	102	2	PH1281	Ig heavy chain V r
511	4	40.0	102	2	PH1282	Ig heavy chain V r
512	4	40.0	102	2	PH1271	Ig heavy chain V r
513	4	40.0	102	2	PH1272	Ig heavy chain V r
514	4	40.0	102	2	PH1273	Ig heavy chain V r
515	4	40.0	102	2	H69841	conserved hypothet
516	4	40.0	102	2	S28361	hypothetical prote
517	4	40.0	102	2	T39407	Ig lambda chain C
518	4	40.0	103	2	B26167	Ig lambda chain C
519	4	40.0	103	2	B60608	myosin heavy chain
520	4	40.0	103	2	I65238	parvalbumin - mous
521	4	40.0	103	2	I51248	ACBP/DBI - duck
522	4	40.0	103	2	G75513	conserved hypothet
523	4	40.0	104	1	GMBO	gastirin precursor
524	4	40.0	104	2	PH1665	Ig heavy chain V r
525	4	40.0	104	2	S69899	Ig heavy chain V r
526	4	40.0	104	2	J00863	hypothetical 11.6k
527	4	40.0	104	2	A83871	probable inner mem
528	4	40.0	104	2	B81787	Ig lambda chain C
529	4	40.0	105	1	L2HU	spdp protein - Str
530	4	40.0	105	2	T36213	Ig lambda chain C
531	4	40.0	105	2	T49655	nodc protein - Bra
532	4	40.0	105	2	S27493	nodc protein - Bra
533	4	40.0	105	2	S44838	KO2D10.3 protein -
534	4	40.0	105	2	D82734	outer membrane pro
535	4	40.0	106	1	CCDV3S	cytochrome c3 - De
536	4	40.0	106	1	K4RBS	Ig kappa-2 chain C
537	4	40.0	106	2	S20774	Ig heavy chain V r
538	4	40.0	106	2	G20907	Ig kappa-B4 chain
539	4	40.0	106	2	G72632	hypothetical prote
540	4	40.0	106	2	S64649	RBL2 protein - yea

249	5	50.0	1098	2	S38100	hypothetical prote	322	4	40.0	48	2	T12940	hypothetical prote
250	5	50.0	1113	2	H84105	hypothetical prote	323	4	40.0	49	2	S39857	ribosomal protein
251	5	50.0	1123	2	S36846	myosin-binding pro	324	4	40.0	49	2	B56448	cofilin-like prote
252	5	50.0	1127	1	E71156	endopeptidase La h	325	4	40.0	49	2	B86888	50S ribosomal prot
253	5	50.0	1132	2	B82538	ribonuclease E Xf2	326	4	40.0	50	2	B60718	phospholipase A2 h
254	5	50.0	1138	2	S24614	myosin-binding pro	327	4	40.0	50	2	F84075	hypothetical prote
255	5	50.0	1150	1	A55289	kinlesin-like prote	328	4	40.0	52	2	A38942	chemoreceptor prot
256	5	50.0	1199	2	T23832	protein-tyrosine k	329	4	40.0	52	2	A69391	ISORP-like protei
257	5	50.0	1210	2	S20969	Na+/Ca2+, K+-exchan	330	4	40.0	52	2	T00141	hypothetical prote
258	5	50.0	1210	2	I39410	AF-4 protein. spli	331	4	40.0	53	2	S66602	aldehyde oxidase (
259	5	50.0	1213	2	A58198	serine/proline-ric	332	4	40.0	53	2	F47395	histone H1 II-1 (c
260	5	50.0	1217	2	T25894	hypothetical prote	333	4	40.0	53	2	B47395	histone H1 II-1 (c
261	5	50.0	1228	2	S59681	probable membrane	334	4	40.0	53	2	C47395	histone H1 II-1 (c
262	5	50.0	1233	2	S56271	hypothetical prote	335	4	40.0	53	2	G47395	histone H1 II-2 (c
263	5	50.0	1245	2	T18211	delta endotoxin -	336	4	40.0	53	2	A72614	hypothetical prote
264	5	50.0	1251	2	A56677	neuronal cell cycl	337	4	40.0	54	2	B36943	chemoreceptor prot
265	5	50.0	1252	2	S02004	I-phosphatidylinos	338	4	40.0	55	2	T03337	gene e30 protein -
266	5	50.0	1265	2	A34163	C-terminal domain-	339	4	40.0	56	2	I56605	hypothetical prote
267	5	50.0	1268	2	T31420	transferrin-like p	340	4	40.0	56	2	B64368	interferon alpha O
268	5	50.0	1274	2	T10729	probable calmoduli	341	4	40.0	58	2	I79344	phosphoprotein Pp3
269	5	50.0	1276	2	T02711	chromosomal protei	342	4	40.0	59	2	B70306	conserved hypotnet
270	5	50.0	1290	2	A55094	acetyl-CoA carboxy	343	4	40.0	61	2	T43672	hypothetical prote
271	5	50.0	1306	2	S42659	nuclear pore membr	344	4	40.0	62	1	CCB017	ubiquinol--cytochr
272	5	50.0	1337	2	A53824	hypothetical prote	345	4	40.0	62	2	T44918	hypothetical prote
273	5	50.0	1341	2	T17285	zyg-9 protein - Ca	346	4	40.0	62	2	T26847	hypothetical prote
274	5	50.0	1415	2	T21244	lin-15B protein -	347	4	40.0	64	2	H82849	hypothetical prote
275	5	50.0	1440	2	T27942	alpha-2-macroglobu	348	4	40.0	65	2	G69542	hypothetical prote
276	5	50.0	1472	2	A26122	salivary agglutini	349	4	40.0	65	2	S35024	hypothetical prote
277	5	50.0	1473	2	A35186	ribosome receptor,	350	4	40.0	66	2	S65971	yyed protein - Bac
278	5	50.0	1534	2	A56734	cell surface anti	351	4	40.0	66	2	PN0644	hypothetical prote
279	5	50.0	1566	2	A43607	hypothetical prote	352	4	40.0	67	2	I47395	histone H1 I-1 (cl
280	5	50.0	1582	2	T15308	acetyl-CoA carboxy	353	4	40.0	67	2	A64321	archaeal histone -
281	5	50.0	1625	2	T02921	adhesin p1 precurs	354	4	40.0	67	2	A64457	archaeal histone -
282	5	50.0	1627	1	IYMAP	adhesin p1, group	355	4	40.0	67	2	D54416	archaeal histone -
283	5	50.0	1635	2	A41480	acetyl-CoA carboxy	356	4	40.0	67	2	D64513	tyrc protein - Bac
284	5	50.0	1685	2	T02750	coracle gene prote	357	4	40.0	68	2	JH0100	small acid-soluble
285	5	50.0	1698	2	T13800	major merizoite su	358	4	40.0	68	2	C83767	lysine kallikrein
286	5	50.0	1785	2	A45546	cag island protein	359	4	40.0	69	2	S28195	synaptoagm II p
287	5	50.0	1819	2	A71928	vitellinogenin I pre	360	4	40.0	69	2	PC7068	conserved hypotnet
288	5	50.0	1912	2	T29088	cag pathogenicity	361	4	40.0	69	2	H69383	50S ribosomal prot
289	5	50.0	1927	2	G64585	myosin heavy chain	362	4	40.0	69	2	C68886	histone H1 I-1 (N-
290	5	50.0	1928	2	S46773	callose synthase c	363	4	40.0	70	2	A47395	RNA-directed RNA p
291	5	50.0	1963	2	T49914	gag, pol and env p	364	4	40.0	70	2	S56769	30S ribosomal prot
292	5	50.0	2272	2	T18572	acetyl-CoA carboxy	365	4	40.0	71	2	C82807	lysine kallikrein
293	5	50.0	2325	2	T02235	zonadhesin - pig	366	4	40.0	72	2	S28196	hypothetical prote
294	5	50.0	2476	2	T34022	hypothetical prote	367	4	40.0	72	2	P00613	hypothetical prote
295	5	50.0	2541	2	T29340	fatty-acid synthas	368	4	40.0	72	2	P00614	hypothetical prote
296	5	50.0	2796	2	UC4743	vacuolar protein s	369	4	40.0	72	2	P00615	conserved hypotnet
297	5	50.0	3131	2	T39553	emulatin synthetas	370	4	40.0	72	2	S30980	hypothetical prote
298	5	50.0	3131	2	S39842	hemocytin - silkw	371	4	40.0	72	2	S75519	hypothetical prote
299	5	50.0	3133	2	S52093	connectin/titin -	372	4	40.0	72	2	E69386	hypothetical prote
300	5	50.0	4162	2	T42633	ALR protein - huma	373	4	40.0	72	2	E69386	hypothetical prote
301	5	50.0	4957	2	T03455	ALR protein - huma	374	4	40.0	72	2	E69386	hypothetical prote
302	5	50.0	5262	2	T03454	protein UNC-89 - C	375	4	40.0	72	2	T48971	conserved hypotnet
303	5	50.0	6642	2	T29757	163K exoantigen -	376	4	40.0	73	2	D69319	hypothetical prote
304	5	50.0	12	2	A33099	conantokin T - con	377	4	40.0	73	2	T44814	hypothetical prote
305	4	40.0	21	1	A35225	calcium-activated	378	4	40.0	73	2	G64443	hypothetical prote
306	4	40.0	22	2	C39800	seed protein ws-6	379	4	40.0	74	2	E69028	hypothetical prote
307	4	40.0	23	2	F61491	potassium channel	380	4	40.0	75	2	T12080	low molecular weigh
308	4	40.0	23	2	P50446	flavodoxin isoform	381	4	40.0	75	2	C85636	hypothetical prote
309	4	40.0	27	2	S55234	30K serine protein	382	4	40.0	75	2	A64842	yyed protein - Esc
310	4	40.0	30	2	B60291	H+-transporting AT	383	4	40.0	76	2	T17673	hypothetical prote
311	4	40.0	32	2	S21245	R-phycocerythrin ga	384	4	40.0	76	2	B83402	hypothetical prote
312	4	40.0	33	2	I22565	primase - Escheric	385	4	40.0	76	2	A75309	hypothetical prote
313	4	40.0	36	2	I57728	hypothetical prote	386	4	40.0	77	2	H69328	nifu protein (nifu
314	4	40.0	38	2	T31479	muonate cycloisom	387	4	40.0	77	2	D85940	hypothetical prote
315	4	40.0	40	2	B54357	hypothetical prote	388	4	40.0	77	2	S00970	krcal protein - pl
316	4	40.0	41	2	B54357	hypothetical prote	389	4	40.0	78	2	S61468	p83/100 protein -
317	4	40.0	41	4	I68622	hypothetical myeli	390	4	40.0	78	2	G71720	hypothetical prote
318	4	40.0	42	4	A47116	trifoliotxin precu	391	4	40.0	78	2	T16325	hypothetical prote
319	4	40.0	44	1	MMVZK4	K4 protein - vacci	392	4	40.0	79	2	T30118	hypothetical prote
320	4	40.0	44	2	S56313	histone H1 II-1 (c	393	4	40.0	80	2	F72306	conserved hypotnet
321	4	40.0	47	2	E47395	hypothetical prote	394	4	40.0	80	2	D75366	hypothetical prote



103	5	50.0	376	2	S04497	surface antigen PA
104	5	50.0	389	2	D84068	RNA helicase BH334
105	5	50.0	391	2	I51097	thyroid hormone re
106	5	50.0	395	2	D83782	aspartate aminotra
107	5	50.0	410	2	D70599	hypothetical prote
108	5	50.0	415	2	T20335	hypothetical prote
109	5	50.0	416	2	G69748	conserved hypothet
110	5	50.0	418	2	F83986	transposer BH2694
111	5	50.0	421	2	T14773	hypothetical prote
112	5	50.0	422	2	G70984	hypothetical prote
113	5	50.0	423	2	S73020	probable pap45 pro
114	5	50.0	428	2	S28468	hypothetical prote
115	5	50.0	435	2	T19687	mannose-1-phosphat
116	5	50.0	440	2	D64616	hypothetical prote
117	5	50.0	443	2	E71335	N-acetylmuramoyl-L
118	5	50.0	445	2	D71977	hypothetical prote
119	5	50.0	445	2	C38537	phosphoglucosamine
120	5	50.0	447	2	T34592	urec protein - Hel
121	5	50.0	451	2	B96774	probable lipoprote
122	5	50.0	452	2	T38962	hypothetical prote
123	5	50.0	453	1	F64623	hypothetical prote
124	5	50.0	453	1	A71891	amidase - Helicoba
125	5	50.0	460	2	T23087	glu-tRNA amidotran
126	5	50.0	461	2	T42513	hypothetical prote
127	5	50.0	465	2	H82345	1D-myo-inositol-tr
128	5	50.0	469	2	H71897	mannose-1-phosphat
129	5	50.0	470	2	T15946	probable probable
130	5	50.0	471	2	JC1403	hypothetical prote
131	5	50.0	473	2	D70405	glutamate--ammonia
132	5	50.0	473	2	S75141	glutamate--trRNA 11
133	5	50.0	476	2	B44997	glutamate--ammonia
134	5	50.0	477	2	A75545	mercoite surface
135	5	50.0	478	2	A32555	amidophosphoribosy
136	5	50.0	482	2	A44997	major mercoite su
137	5	50.0	484	2	D65230	mercoite surface
138	5	50.0	484	2	A86116	hypothetical 52.9
139	5	50.0	490	2	PQ0164	hypothetical prote
140	5	50.0	492	2	T43859	sucrose phosphoryl
141	5	50.0	493	2	T12044	uroprophyrinogen I
142	5	50.0	496	2	T30976	hypothetical legumain
143	5	50.0	498	2	D64878	hypothetical prote
144	5	50.0	498	2	G85755	probable glutamate
145	5	50.0	501	2	T27513	glutamate--ammonia
146	5	50.0	504	2	G02474	hypothetical prote
147	5	50.0	505	2	T01667	interferon regulat
148	5	50.0	505	2	B64560	gag polyprotein -
149	5	50.0	507	2	G01614	poly E-rich protei
150	5	50.0	509	1	S04346	zinc finger protei
151	5	50.0	516	2	T09364	steroid 17alpha-mo
152	5	50.0	519	2	S23796	cytochrome P450 ho
153	5	50.0	519	2	T45038	modifier-3 protein
154	5	50.0	520	2	E96795	hypothetical prote
155	5	50.0	530	2	D83281	unknown protein F2
156	5	50.0	532	2	S40963	probable MFS trans
157	5	50.0	537	2	D86520	hypothetical prote
158	5	50.0	537	2	A72103	CMP synthetase (lm
159	5	50.0	537	2	E81568	CTP synthetase - C
160	5	50.0	539	2	I49065	CRP synthase CP052
161	5	50.0	539	2	T02508	lymphoid-restrict
162	5	50.0	542	2	T29707	hypothetical prote
163	5	50.0	543	2	T00513	hypothetical prote
164	5	50.0	547	2	B56573	cytochrome P450 ho
165	5	50.0	550	2	G86704	nuclear pore compl
166	5	50.0	553	2	C75318	conserved hypothet
167	5	50.0	557	2	A70480	hypothetical prote
168	5	50.0	557	2	B83962	carbamoyl-phosphat
169	5	50.0	557	2	A00629	hypothetical prote
170	5	50.0	558	2	B86016	protein-tyrosine k
171	5	50.0	558	2	S75104	probable enzyme z4
172	5	50.0	562	2	G84221	hypothetical prote
173	5	50.0	566	2	C70482	NADH dehydrogenase
174	5	50.0	567	2	B86481	repair excision nu
175	5	50.0	572	1	FOHYTH	hypothetical prote
						retrovirus-related
176	5	50.0	573	2	T21653	hypothetical prote
177	5	50.0	573	2	S66710	probable membrane
178	5	50.0	574	1	S76132	protein kinase pkn
179	5	50.0	578	2	T33916	hypothetical prote
180	5	50.0	580	2	C96597	hypothetical prote
181	5	50.0	583	2	S01301	Rubisco subunit b1
182	5	50.0	583	2	T30131	lamini L-III - Afri
183	5	50.0	603	2	S06059	hypothetical prote
184	5	50.0	604	2	S25203	gene NDI Intron 4
185	5	50.0	604	2	S25203	srmr protein - str
186	5	50.0	612	2	G71972	threonine--trNA 11
187	5	50.0	612	2	C64535	threonine--trNA 11
188	5	50.0	615	2	T37981	probable exocyst c
189	5	50.0	637	2	S66953	hypothetical prote
190	5	50.0	641	2	T03095	homeoprotein Sail
191	5	50.0	643	1	DJEC3G	DNA-directed DNA p
192	5	50.0	646	2	G85544	hypothetical prote
193	5	50.0	646	2	G69871	hypothetical prote
194	5	50.0	651	2	T21175	hypothetical prote
195	5	50.0	654	2	HMIVC8	hemagglutinin prec
196	5	50.0	655	2	S07412	hemagglutinin-este
197	5	50.0	662	2	T20570	hypothetical prote
198	5	50.0	668	2	C71365	probable excinucle
199	5	50.0	668	2	H86729	exclnuclease ABC s
200	5	50.0	670	2	T34548	hypothetical prote
201	5	50.0	691	2	H81313	ATP-dependent DNA
202	5	50.0	700	2	T13577	hypothetical prote
203	5	50.0	701	1	S33709	DHR39-short protel
204	5	50.0	707	2	JC2218	procollagen C-endo
205	5	50.0	713	2	A28706	calpastatin, cardl
206	5	50.0	730	1	BMH01	procollagen C-endo
207	5	50.0	739	2	T29407	hypothetical prote
208	5	50.0	746	2	T19287	hypothetical prote
209	5	50.0	752	1	H70628	hypothetical prote
210	5	50.0	752	1	S51866	probable pknrg prot
211	5	50.0	763	2	T44735	hPR1 protein - yea
212	5	50.0	771	2	T26176	serine/threonine-s
213	5	50.0	775	2	E83400	hypothetical prote
214	5	50.0	792	2	S20554	pyrroloquinoline q
215	5	50.0	792	2	E85778	pyruvate,water dlk
216	5	50.0	806	2	H70647	phosphoenolpyruvat
217	5	50.0	808	1	S33708	probable steroid/th
218	5	50.0	811	2	PN0689	nuclear connectin 1
219	5	50.0	815	2	H96494	connectin 1 - chic
220	5	50.0	816	2	T17257	protein F7F22.2 [1
221	5	50.0	817	2	S51342	hypothetical prote
222	5	50.0	822	2	G69010	verpoinin - yeast
223	5	50.0	823	1	A58788	conserved hypothet
224	5	50.0	871	2	H81430	procollagen C-endo
225	5	50.0	871	2	D86355	translacion Initia
226	5	50.0	887	2	E81847	protein T16E15.12
227	5	50.0	887	2	E81094	pyruvate dehydroge
228	5	50.0	899	2	S51341	pyruvate dehydroge
229	5	50.0	912	2	H71931	SGDI protein - yea
230	5	50.0	919	2	C86578	hypothetical prote
231	5	50.0	925	2	F72045	topc/gen secretion
232	5	50.0	925	2	T33732	type III secretion
233	5	50.0	958	2	S47179	probable excinucle
234	5	50.0	969	2	A75634	hypothetical prote
235	5	50.0	980	2	S71090	McrB-related prote
236	5	50.0	982	2	T43659	peroxisome biogene
237	5	50.0	986	2	B58788	DNA mismatch repai
238	5	50.0	990	2	T12678	procollagen C-endo
239	5	50.0	991	2	T149540	hypothetical prote
240	5	50.0	998	2	H75005	procollagen C-endo
241	5	50.0	1000	2	B72385	ATP-dependent prot
242	5	50.0	1001	2	T29105	hypothetical prote
243	5	50.0	1002	2	S70292	myosin-B - Toxopla
244	5	50.0	1021	2	I39207	leukocyte surface
245	5	50.0	1025	2	E86355	hypothetical prote
246	5	50.0	1070	2	T34385	hypothetical prote
247	5	50.0	1071	2	T04926	starch synthase ho
248	5	50.0	1073	2	T01955	hypothetical prote
						carbamoyl-phosphat

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model.

Run on: August 15, 2001, 12:34:44 ; Search time 33.87 Seconds  
(without alignments)  
22.490 Million cell updates/sec

Title: US-09-372-036-31  
Perfect score: 10  
Sequence: 1 PVAPTEVKK 10

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR\_68:\*

1: pir1:\*\n2: pir2:\*\n3: pir3:\*\n4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	100.0	484	2 A14187	protein p60 precur
2	7	70.0	2 D81259		probable tonb tran
3	6	60.0	102	2 F83231	hypothetical prote
4	6	60.0	116	2 T37186	probable regulator
5	6	60.0	456	2 H85574	hypothetical prote
6	6	60.0	617	2 S42719	actin-binding prot
7	6	60.0	748	2 H84913	probable ATP-depen
8	6	60.0	985	1 VCLJSP	env polyprotein
9	6	60.0	1133	2 T30302	P-type ATPase - Te
10	6	60.0	1230	2 E64664	outer membrane pro
11	5	50.0	75	2 E75264	hypothetical prote
12	5	50.0	82	2 T29802	hypothetical prote
13	5	50.0	102	2 B56581	major basic nuclea
14	5	50.0	103	1 KARB	Ig kappa-B4 chain
15	5	50.0	104	2 E53275	Ig kappa-1 chain C
16	5	50.0	112	2 T14788	hypothetical prote
17	5	50.0	114	2 T08411	hypothetical prote
18	5	50.0	124	2 B86557	Dnak suppressor [i
19	5	50.0	124	2 A71518	probable dnak supp
20	5	50.0	124	2 G72066	dna suppressor -
21	5	50.0	124	2 T24876	hypothetical prote
22	5	50.0	131	2 D66684	phophage p11 prote
23	5	50.0	132	2 B81600	conserved hypotnet
24	5	50.0	133	2 T50083	probable cytidine
25	5	50.0	134	2 A81676	conserved hypotnet
26	5	50.0	141	2 E72580	hypothetical prote
27	5	50.0	147	2 E72384	conserved hypotnet
28	5	50.0	147	2 T33327	hypothetical prote
29	5	50.0	155	2 S35162	SRH-21 protein - p

30	5	50.0	157	2 T25638	hypothetical prote
31	5	50.0	161	2 C96568	hypothetical prote
32	5	50.0	174	2 S07291	repressor protein
33	5	50.0	175	2 D66699	hypothetical prote
34	5	50.0	179	2 G84789	hypothetical prote
35	5	50.0	185	2 H86513	elongation factor
36	5	50.0	185	2 G72110	translacion elonga
37	5	50.0	185	2 D96572	protein f12M16.6 f
38	5	50.0	187	2 JC4806	core protein G - p
39	5	50.0	190	2 H75162	anthranilate synth
40	5	50.0	196	2 A41050	repressor protein
41	5	50.0	199	2 F66832	hypothetical prote
42	5	50.0	200	2 F64217	ribosomal protein
43	5	50.0	201	1 IMPPSR	immunity protein -
44	5	50.0	206	2 A71326	probable V-type AT
45	5	50.0	209	2 S04827	gene 40A protein -
46	5	50.0	211	2 T22122	hypothetical prote
47	5	50.0	217	2 S10212	late 33k protein -
48	5	50.0	217	2 T37859	probable transcrip
49	5	50.0	218	2 C68847	transcription regu
50	5	50.0	221	2 H86860	protein serine/chr
51	5	50.0	227	2 T12797	immunity protein d
52	5	50.0	228	2 H71214	hypothetical prote
53	5	50.0	239	2 A27207	tissue kallikrein
54	5	50.0	241	2 S75154	urase accessory p
55	5	50.0	246	2 I40763	hypothetical prote
56	5	50.0	246	2 E81373	hypothetical prote
57	5	50.0	250	2 S62838	hypothetical prote
58	5	50.0	251	2 T45707	CHLOROPHYLL A-B B
59	5	50.0	252	2 A86449	hypothetical prote
60	5	50.0	253	2 E71198	hypothetical prote
61	5	50.0	253	2 A75159	protein-1-isoaspar
62	5	50.0	254	2 T23323	hypothetical prote
63	5	50.0	260	2 T18909	hypothetical prote
64	5	50.0	263	2 C69887	hypothetical prote
65	5	50.0	271	2 T16590	hypothetical prote
66	5	50.0	276	2 S75249	esterase s110992 -
67	5	50.0	276	2 T29894	hypothetical prote
68	5	50.0	287	2 E75159	hypothetical prote
69	5	50.0	288	2 A81009	hypothetical prote
70	5	50.0	289	2 C70400	ferredoxin oxidore
71	5	50.0	291	2 S42420	replication protei
72	5	50.0	293	2 T09758	myb-related protei
73	5	50.0	293	2 A82690	phosphatidylserine
74	5	50.0	296	2 S46018	probable membrane
75	5	50.0	296	2 G84747	AT-hook DNA-bindin
76	5	50.0	305	2 JN0518	acetyltransferase (E
77	5	50.0	305	2 F83966	dhlydrocorolate deh
78	5	50.0	307	1 S75769	hypothetical prote
79	5	50.0	308	2 E75409	probable antibioci
80	5	50.0	309	1 A25776	urate oxidase (EC
81	5	50.0	315	2 T25171	hypothetical prote
82	5	50.0	321	2 D70070	transcription regu
83	5	50.0	327	2 JC4195	36k hydrophobic nu
84	5	50.0	327	2 H82554	conserved hypotnet
85	5	50.0	335	2 G75309	lybA protein - Del
86	5	50.0	335	2 T37262	probable tyrosine
87	5	50.0	336	2 B38274	Y box-binding prot
88	5	50.0	337	1 T46742	ornithine carbamoy
89	5	50.0	340	2 D86837	ornithine carbamoy
90	5	50.0	342	2 T45456	probable membrane
91	5	50.0	343	2 T14245	MDH dehydrogenase
92	5	50.0	343	2 H82171	probable multidrug
93	5	50.0	346	2 S56665	tryptophan synthas
94	5	50.0	355	2 C70457	hypothetical prote
95	5	50.0	360	2 S32695	hypothetical prote
96	5	50.0	360	2 T26037	Wnt-2 protein - Ca
97	5	50.0	366	1 MNXRSD	sigma NS protein -
98	5	50.0	366	1 MNXRST	sigma NS protein -
99	5	50.0	366	1 MNXRRT	sigma NS protein -
100	5	50.0	367	2 C27401	sigma NS protein -
101	5	50.0	367	2 S46336	env polyprotein -
102	5	50.0	368	2 S06058	MDH dehydrogenase

***This Page Blank (uspto)***

PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 5; DB 21; Length 217;  
Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
11111  
Db 160 pvapt 164

Search completed: August 15, 2001, 12:35:49  
Job time: 174 sec

PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
  
PR 13-JUL-1999; 99US-0143342.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
PR 27-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-015138.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0156559.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.

```

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      50.0%; Score 5; DB 21; Length 196;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 PVAPT 5
    |||||
Db 32 pvapt 36

```

```

RESULT 49
AAW89834

```

```

ID AAW89834 standard; Protein: 201 AA.

```

```

AC AAW89834;

```

```

DT 18-FEB-1999 (first entry)

```

```

XX Protein encoded by clone c5.
XX

```

```

KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
XX
OS Helicobacter pylori.
XX
PN WO9849314-A2.
XX
PD 05-NOV-1998.
XX
PE 27-APR-1998; 98WO-US08487.
XX
PR 14-OCT-1997; 97US-0061958.
PR 25-APR-1997; 97US-0045107.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Chow TP, Fry KE, Lim MY, McAtee CP;
XX WPI; 1999-009433/01.
XX DR N-PSDB; AAV90768.
XX
PT New Helicobacter pylori antigens and related nucleic acid sequences
PT - useful in serological diagnosis and protective vaccines, providing
PT long-lasting immune response
XX
PS Claim 15; Page 180-181; 402pp; English.
XX
CC The present sequence represents a Helicobacter pylori antigenic protein
CC that is characterised by immunoreactivity with H. pylori-positive
CC antisera. The proteins are highly immunogenic and induce a long-lasting
CC immune response that persists even after antimicrobial treatment. In
CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are
CC highly sensitive and specific. The specification also describes 69
CC previously unrecognised immunogenic cluster families. H. pylori antigens
CC are used to detect H. pylori-specific antibodies, for diagnosing
CC infection or to confirm eradication of infection, and in vaccines to
CC protect against H. pylori infection and related diseases (gastritis,
CC peptic ulcer, gastric adenocarcinoma/lymphoma).
XX
SQ Sequence 201 AA;

```

```

Query Match      50.0%; Score 5; DB 20; Length 201;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 6 QEVKK 10
    .|||||
Db 139 gevkk 143

```

```

RESULT 50

```

```

ID AAG04638 standard; Protein: 217 AA.

```

```

AC AAG04638;

```

```

DT 17-OCT-2000 (first entry)

```

```

XX Arabidopsis thaliana protein fragment SEQ ID NO: 738.

```

```

XX Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

```

```

XX Arabidopsis thaliana.

```

```

OS EP1033405-A2.

```

```

XX 06-SEP-2000.

```

```

XX 25-FEB-2000; 2000EP-0301439.
XX

```

XX AAC59940;  
AC  
XX 18-OCT-2000 (first entry)  
DT  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 77587.  
DE  
XX Protein identification: signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX EPI033405-A2.  
PN  
XX 06-SEP-2000.  
PD  
XX 25-FEB-2000; 2000EP-0301439.  
PE  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
XX 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
XX 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
XX 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
XX 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
XX 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
XX 24-MAY-1999; 99US-0135533.  
PR 25-MAY-1999; 99US-0135629.  
XX 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
XX 01-JUN-1999; 99US-0136782.  
PR 03-JUN-1999; 99US-0137222.  
XX 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137502.  
XX 08-JUN-1999; 99US-0137724.  
PR 10-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138340.  
PR 14-JUN-1999; 99US-0138847.  
XX 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
XX 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
XX 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
XX 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
XX 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.  
XX 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
XX 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
XX 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
XX 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
XX 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
XX 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
XX 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
XX 06-JUL-1999; 99US-0142290.  
PR 08-JUL-1999; 99US-0142803.  
XX 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
XX 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
XX 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
XX 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
XX 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
XX 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
XX 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
XX 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
XX 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
XX 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
XX 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
XX 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
XX 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145224.  
XX 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
XX 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
XX 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
XX 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
XX 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
XX 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147192.  
XX 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
XX 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
XX 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
XX 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
XX 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
XX 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
XX 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
XX 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
XX 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.

PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144003.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161922.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161992.
PR	04-AUG-1999;	99US-0147204.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.	Query Match		
PR	05-AUG-1999;	99US-0147260.	Best Local Similarity		
PR	06-AUG-1999;	99US-0147303.	Matches		
PR	06-AUG-1999;	99US-0147416.	5: Conservative		
PR	09-AUG-1999;	99US-0147493.	0: Mismatches		
PR	09-AUG-1999;	99US-0147935.	0: Indels		
PR	10-AUG-1999;	99US-0148171.	0: Gaps		
PR	11-AUG-1999;	99US-0148319.	0: Indels		
PR	12-AUG-1999;	99US-0148341.	0: Gaps		
PR	13-AUG-1999;	99US-0148565.	0: Indels		
PR	13-AUG-1999;	99US-0148684.	0: Gaps		
PR	16-AUG-1999;	99US-0149368.	0: Indels		

Query Match

Best Local Similarity 50.0%; Score 5; DB 21; Length 196;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVAPT 5

DB 32 pvapt 36

RESULT 48  
AAC59940  
ID AAC59940 standard; Protein: 196 AA.



CC useful for detecting the presence of HCV in a sample, the primers  
CC are also useful for HCV genotyping. Proteins encoded by the cDNAs  
CC can be used in vaccines for immunising against HCV infection. The  
CC proteins may also be used to detect antibodies against HCV in serum,  
CC saliva, lymphocytes or other mononuclear cells. The antibodies may be  
CC used in the prevention of HCV infection.

XX Sequence 192 AA;

Query Match 50.0%; Score 5; DB 17; Length 192;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PVAPT 5

Db 50 pvapt 54

RESULT 46

AAR33986

ID AAR33986 standard; Protein; 193 AA.

XX AAR33986;

DT 26-JUL-1993 (first entry)

XX HCT27 E1 protein.

KW Polymerase chain reaction; PCR; amplify; primer; hepatitis C virus;

KW HCV; asymptomatic; chronically infected; epitope; viral isolate;

KW domain; immunological; cross-reactive; envelope protein.

XX Synthetic.

XX WO9306126-A.

XX 01-APR-1993.

XX 11-SEP-1992; 92MO-US07683.

XX 13-SEP-1991; 91US-0759575.

XX (CHIR ) CHIRON CORP.

XX Houghton M, Weiner AJ;

XX WPI; 1993-117468/14.

XX Immuno-reactive hepatitis C virus polypeptide compns. - contg.

XX at least 2 sequences from the first variable domain of distinct

XX HCV isolates

XX Disclosure; Fig 2; 106pp; English.

XX The sequences given in AAR33982-91 represent a portion of the E1 protein

XX encoded by group I and group II HCV isolates, from amino acid 192-383.

XX E1 is a viral envelope protein and is of immunogenic interest. E1

XX contains a variable domain between amino acids 215-255. This is an

XX important immunoreactive domain. It has been discovered that a number

XX of important HCV epitopes vary among viral isolates and that these

XX epitopes can be mapped to specific domains. This meant that immuno-

XX logically cross-reactive polypeptides which focus on variable rather

XX than constant domains can be produced.

XX Sequence 193 AA;

Query Match

Best Local Similarity 100.0%; Score 5; DB 14; Length 193;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PVAPT 5

Db 50 pvapt 54

RESULT 47

ID AAG55984 standard; Protein; 196 AA.

XX AAG55984;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 71881.

DE Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 28-APR-1999; 99US-0130891.

XX 30-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 07-MAY-1999; 99US-0132487.

XX 11-MAY-1999; 99US-0132863.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

XX 14-MAY-1999; 99US-0134321.

XX 18-MAY-1999; 99US-0134370.

XX 18-MAY-1999; 99US-0134768.

XX 19-MAY-1999; 99US-0134941.

XX 20-MAY-1999; 99US-0135124.

XX 21-MAY-1999; 99US-0135353.

XX 24-MAY-1999; 99US-0135629.

XX 25-MAY-1999; 99US-0136021.

XX 27-MAY-1999; 99US-0136392.

XX 28-MAY-1999; 99US-0136782.

XX 01-JUN-1999; 99US-0137222.

XX 03-JUN-1999; 99US-0137528.

XX 04-JUN-1999; 99US-0137502.

XX 07-JUN-1999; 99US-0137724.

XX 08-JUN-1999; 99US-0138094.

XX 10-JUN-1999; 99US-0138540.

XX 10-JUN-1999; 99US-0138847.

XX 14-JUN-1999; 99US-0139119.

XX 16-JUN-1999; 99US-0139452.

XX 17-JUN-1999; 99US-0139453.

XX 18-JUN-1999; 99US-0139454.

```

XX 13-SEP-1999 (first entry)
DT Chlamydia pneumoniae transmembrane protein sequence.
DE
XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
KW vaccine; neutralising epitope.
XX Chlamydia pneumoniae.
OS
XX MO9927105-A2.
PN
XX 03-JUN-1999.
PD
XX 20-NOV-1998; 98WO-1B01890.
PE
XX 04-NOV-1998; 98US-0107078.
PR 21-NOV-1997; 97FR-0014673.
XX
XX (GEST ) GENSET.
PA
XX Grifvals R;
PI
XX WPI; 1999-357842/30.
DR
XX Genome sequence of Chlamydia pneumoniae
PT
XX Page 759-760; Disclosure; 1912pp; English.
PS
XX AAY34584-Y35879 represent the proteins encoded by all the open reading
XX frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
CC C. pneumoniae causes respiratory disease such as pneumonia and
CC bronchitis and is thought to be a contributing factor in heart
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
CC
XX
XX Sequence 185 AA;
SQ

```

Query Match 50.0%; Score 5; DB 20; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 5 TOEVK 9
   |||||
Db 57 tqevk 61

```

RESULT 44  
 AAR69639  
 ID AAR69639 standard; Protein; 192 AA.  
 XX  
 XX AAR69639;  
 AC  
 XX 14-SEP-1995 (first entry)  
 DT  
 XX Hepatitis C virus envelope 1 protein isolate S18.  
 DE  
 XX Hepatitis C virus; HCV; non-A non-B; envelope 1 gene; isolate S18;  
 KW diagnosis; vaccines; antibodies; antisera; gene inhibition.  
 XX  
 XX Hepatitis C virus.  
 OS  
 XX MO9501442-A.  
 PN  
 XX 12-JAN-1995.  
 PD  
 XX

```

PF 28-JUN-1994; 94WO-US07320.
XX
XX 29-JUN-1993; 93US-0086428.
PR
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX Burk J, Miller RH, Purcell RH;
PI
XX WPI; 1995-061006/08.
DR
XX N-PSDB; AA083850.
DR
XX Envelope 1 cDNAs of 51 hepatitis C virus isolates - and derived
PT oligo-nucleotide(s), peptide(s) and proteins, used in diagnosis
PT and in vaccines
PT
XX Claim 3; Page 84; 186pp; English.
PS
XX AA083850 encodes AAR69639 hepatitis C virus (HCV) envelope 1 (E1)
CC protein isolate S18, both can be used for the diagnosis of HCV
CC infection, and in the prodn. of anti-HCV vaccines; antibodies
CC and antisera. The cDNA may also be used to inhibit the expression
CC of the HCV E1 gene.
CC
XX
XX Sequence 192 AA;
SQ

```

Query Match 50.0%; Score 5; DB 16; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 PVAPT 5
   |||||
Db 50 pvapt 54

```

RESULT 45  
 AAR89510  
 ID AAR89510 standard; Protein; 192 AA.  
 XX  
 XX AAR89510;  
 AC  
 XX 30-SEP-1996 (first entry)  
 DT  
 XX Hepatitis C virus isolate S18 envelope 1 protein.  
 DE  
 XX HCV; E1; envelope 1; core protein; HCV genotyping; antibody; vaccine;  
 KW hepatitis.  
 KW  
 XX Hepatitis C virus.  
 OS  
 XX MO9605315-A2.  
 PN  
 XX 22-FEB-1996.  
 PD  
 XX 15-AUG-1995; 95WO-US10398.  
 PE  
 XX 15-AUG-1994; 94US-0290665.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (USSH ) US SEC DEPT HEALTH.  
 XX  
 XX Burk J, Miller RH, Purcell RH;  
 PI  
 XX WPI; 1996-139709/14.  
 DR  
 XX N-PSDB; AAT16564.  
 DR  
 XX DNA and amino acid sequence of HCV envelope 1 and core proteins -  
 PT used to determine HCV genotype and as vaccines against HCV infection  
 PT  
 XX Claim 2; Page 114; 340pp; English.  
 PS  
 XX AAR89505-R89555 are HCV E1 (envelope-1) proteins derived from 51 HCV  
 CC isolates. Isolated cDNA sequences are used for the prodn. of primers  
 CC

PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161982.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match	50.0%;	Score 5;	DB 21;	Length 180;
Best Local Similarity	100.0%;	Pred. No. 3.1e+02;		
Matches	5;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

```
Qy      1 PVAPT 5
         |||||
Db      16 PVAPT 20
```

RESULT	42
AAB25268	
ID	AAB25268 standard; Protein; 181 AA.
XX	
AC	AAB25268;

Eucalyptus grandis cell signalling involved protein SEQ ID NO:587

Euclalyptus grandis; Pinus radiata; Monterey pine; plant; modification; plant cell signaling; modulation; transgenic plant; pathogen; growth; environmental change; development; cell proliferation; differentiation; elongation; survival; disease resistance; nutrient metabolism.

*Eucalyptus grandis*

WO200042171-A1

20-JUL-2000.

11-JAN-2000; 2000WO-US00724.

12-JAN-1999; 99US-0228986

01-NOV-1999; 99US-0162866.

(GENE-) GENESIS RES & DEV CORP LTD.

Strabala TJ, Nieuwenhuizen NJ;

WPI; 2000-476052/41.

Isolated polynucleotide encoding a polypeptide involved in cell signaling used for generating transgenic plants with modified responses to external signals -

Claim 3; Page 263; 527pp; English.

AAV9263 to AAV79736 and AAB25100 to AAB25570 represent polynucleotide and protein sequences isolated from eucalyptus (*Eucalyptus grandis*) or pine (*Pinus radiata* also known as Monterey pine). The protein sequences are involved in cell signalling. The polynucleotide and protein sequences can be used to modify the response of plant cells to external signals e.g. environmental changes or pathogens during the growth and development of a plant. They can be used to modify cell proliferation, differentiation, elongation and survival, resistance to disease and nutrient metabolism. Examples of modifications which can be produced are altered fruit ripening and senescence of leaves and flowers e.g. to delay senescence and prolong the life of cut flowers or enhance senescence of reproductive organs to engineer sterile plants. Other modifications can be used to delay senescence in selected cell types or organs providing fruit and vegetables which have a longer shelf life between harvest and consumption, or to decrease branching frequency in forest tree species giving long stretches of valuable knot-free clear wood which can be used in solid timber furniture and veneers.

Sequence 181 AA;

Query Match	50.0%;	Score 5;	DB 21;	Length 181;
Best Local Similarity	100.0%;	Pred. No. 3.1e+02;		
Matches	5;	Conservative 0;	Mismatches 0;	Indels 0;
			Gaps	0

```
QY      4 PTQEV 8
         |||||
Db      162 ptgev 166
```

RESULT	43
AAV34781	
ID	AAV34781 standard; Protein; 185 AA
XX	
AC	AAV34781;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 pVApT 5  
| | | | |  
Db 16 pVapT 20

RESULT 41  
ID AAG59941 standard; Protein: 160 AA.  
XX AAG59941;  
AC AAG59941;  
XX AAG59941;  
DT 18-OCT-2000 (first entry)  
XX 18-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 77586.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 15-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 21-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0145986.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147028.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.

PR 04-JUN-1999: 99US-0137502.  
PR 07-JUN-1999: 99US-0137724.  
PR 08-JUN-1999: 99US-0138094.  
PR 10-JUN-1999: 99US-0138840.  
PR 10-JUN-1999: 99US-0138847.  
PR 14-JUN-1999: 99US-0139119.  
PR 16-JUN-1999: 99US-0139452.  
PR 16-JUN-1999: 99US-0139453.  
PR 17-JUN-1999: 99US-0139454.  
PR 18-JUN-1999: 99US-0139454.  
PR 18-JUN-1999: 99US-0139455.  
PR 18-JUN-1999: 99US-0139456.  
PR 18-JUN-1999: 99US-0139457.  
PR 18-JUN-1999: 99US-0139458.  
PR 18-JUN-1999: 99US-0139459.  
PR 18-JUN-1999: 99US-0139460.  
PR 18-JUN-1999: 99US-0139461.  
PR 18-JUN-1999: 99US-0139462.  
PR 18-JUN-1999: 99US-0139463.  
PR 18-JUN-1999: 99US-0139750.  
PR 18-JUN-1999: 99US-0139753.  
PR 21-JUN-1999: 99US-0139817.  
PR 22-JUN-1999: 99US-0139899.  
PR 23-JUN-1999: 99US-0140353.  
PR 23-JUN-1999: 99US-0140354.  
PR 24-JUN-1999: 99US-0140695.  
PR 28-JUN-1999: 99US-0140823.  
PR 28-JUN-1999: 99US-0140823.  
PR 30-JUN-1999: 99US-0141287.  
PR 30-JUN-1999: 99US-0141287.  
PR 01-JUL-1999: 99US-0141842.  
PR 01-JUL-1999: 99US-0142154.  
PR 02-JUL-1999: 99US-0142055.  
PR 06-JUL-1999: 99US-0142390.  
PR 08-JUL-1999: 99US-0142803.  
PR 09-JUL-1999: 99US-0142820.  
PR 12-JUL-1999: 99US-0142977.  
PR 13-JUL-1999: 99US-0143542.  
PR 14-JUL-1999: 99US-0143624.  
PR 15-JUL-1999: 99US-0144005.  
PR 16-JUL-1999: 99US-0144085.  
PR 16-JUL-1999: 99US-0144086.  
PR 19-JUL-1999: 99US-0144325.  
PR 19-JUL-1999: 99US-0144331.  
PR 19-JUL-1999: 99US-0144332.  
PR 19-JUL-1999: 99US-0144333.  
PR 19-JUL-1999: 99US-0144334.  
PR 19-JUL-1999: 99US-0144335.  
PR 20-JUL-1999: 99US-0144352.  
PR 20-JUL-1999: 99US-0144632.  
PR 20-JUL-1999: 99US-0144684.  
PR 21-JUL-1999: 99US-0144814.  
PR 21-JUL-1999: 99US-0145086.  
PR 21-JUL-1999: 99US-0145088.  
PR 22-JUL-1999: 99US-0145085.  
PR 22-JUL-1999: 99US-0145087.  
PR 22-JUL-1999: 99US-0145089.  
PR 22-JUL-1999: 99US-0145192.  
PR 23-JUL-1999: 99US-0145145.  
PR 23-JUL-1999: 99US-0145218.  
PR 23-JUL-1999: 99US-0145224.  
PR 26-JUL-1999: 99US-0145226.  
PR 27-JUL-1999: 99US-0145913.  
PR 27-JUL-1999: 99US-0145918.  
PR 27-JUL-1999: 99US-0145919.  
PR 28-JUL-1999: 99US-0145951.  
PR 02-AUG-1999: 99US-0146386.  
PR 02-AUG-1999: 99US-0146388.  
PR 02-AUG-1999: 99US-0146389.  
PR 03-AUG-1999: 99US-0147038.  
PR 04-AUG-1999: 99US-0147204.  
PR 04-AUG-1999: 99US-0147302.  
PR 05-AUG-1999: 99US-0147192.  
PR 05-AUG-1999: 99US-0147260.

PR 06-AUG-1999: 99US-0147303.  
PR 06-AUG-1999: 99US-0147416.  
PR 09-AUG-1999: 99US-0147493.  
PR 09-AUG-1999: 99US-0147935.  
PR 10-AUG-1999: 99US-0148171.  
PR 11-AUG-1999: 99US-0148319.  
PR 12-AUG-1999: 99US-0148341.  
PR 13-AUG-1999: 99US-0148565.  
PR 13-AUG-1999: 99US-0148684.  
PR 16-AUG-1999: 99US-0149368.  
PR 17-AUG-1999: 99US-0149375.  
PR 18-AUG-1999: 99US-0149426.  
PR 20-AUG-1999: 99US-0149722.  
PR 20-AUG-1999: 99US-0149723.  
PR 20-AUG-1999: 99US-0149929.  
PR 23-AUG-1999: 99US-0149902.  
PR 23-AUG-1999: 99US-0149930.  
PR 23-AUG-1999: 99US-0149930.  
PR 25-AUG-1999: 99US-0150566.  
PR 26-AUG-1999: 99US-0150884.  
PR 27-AUG-1999: 99US-0151065.  
PR 27-AUG-1999: 99US-0151066.  
PR 27-AUG-1999: 99US-0151080.  
PR 30-AUG-1999: 99US-0151303.  
PR 31-AUG-1999: 99US-0151438.  
PR 01-SEP-1999: 99US-0151330.  
PR 07-SEP-1999: 99US-0152363.  
PR 10-SEP-1999: 99US-0153070.  
PR 13-SEP-1999: 99US-0153758.  
PR 13-SEP-1999: 99US-0154018.  
PR 16-SEP-1999: 99US-0154039.  
PR 20-SEP-1999: 99US-0154779.  
PR 22-SEP-1999: 99US-0155139.  
PR 23-SEP-1999: 99US-0155486.  
PR 24-SEP-1999: 99US-0155659.  
PR 28-SEP-1999: 99US-0156458.  
PR 29-SEP-1999: 99US-0156596.  
PR 04-OCT-1999: 99US-0157117.  
PR 05-OCT-1999: 99US-0157753.  
PR 06-OCT-1999: 99US-0157865.  
PR 07-OCT-1999: 99US-0158029.  
PR 08-OCT-1999: 99US-0158232.  
PR 12-OCT-1999: 99US-0158369.  
PR 13-OCT-1999: 99US-0159293.  
PR 13-OCT-1999: 99US-0159294.  
PR 13-OCT-1999: 99US-0159295.  
PR 14-OCT-1999: 99US-0159329.  
PR 14-OCT-1999: 99US-0159330.  
PR 14-OCT-1999: 99US-0159331.  
PR 14-OCT-1999: 99US-0159637.  
PR 14-OCT-1999: 99US-0159638.  
PR 18-OCT-1999: 99US-0159584.  
PR 21-OCT-1999: 99US-0160741.  
PR 21-OCT-1999: 99US-0160767.  
PR 21-OCT-1999: 99US-0160768.  
PR 21-OCT-1999: 99US-0160770.  
PR 21-OCT-1999: 99US-0160814.  
PR 21-OCT-1999: 99US-0160815.  
PR 22-OCT-1999: 99US-0160980.  
PR 22-OCT-1999: 99US-0160981.  
PR 22-OCT-1999: 99US-0160989.  
PR 25-OCT-1999: 99US-0161404.  
PR 25-OCT-1999: 99US-0161405.  
PR 25-OCT-1999: 99US-0161406.  
PR 26-OCT-1999: 99US-0161359.  
PR 26-OCT-1999: 99US-0161360.  
PR 26-OCT-1999: 99US-0161361.  
PR 28-OCT-1999: 99US-0161920.  
PR 28-OCT-1999: 99US-0161993.  
PR 29-OCT-1999: 99US-0162142.

Query Match 50.0%; Score 5; DB 21; Length 180;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;

PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155658.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 50.0%; Score 5; DB 21; Length 179;  
Best local similarity 100.0%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 PVAFT 5  
Db 15 pvaft 19  
  
RESULT 40  
AGS5985  
ID AGS5985 standard; Protein: 180 AA.  
XX AC AGS5985;  
XX DT 18-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 71882.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hydridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.

PR 22-OCT-1999; 990S-0160989.  
PR 25-OCT-1999; 990S-0161404.  
PR 25-OCT-1999; 990S-0161405.  
PR 25-OCT-1999; 990S-0161406.  
PR 26-OCT-1999; 990S-0161359.  
PR 26-OCT-1999; 990S-0161360.  
PR 26-OCT-1999; 990S-0161361.  
PR 28-OCT-1999; 990S-0161920.  
PR 28-OCT-1999; 990S-0161992.  
PR 28-OCT-1999; 990S-0161993.  
PR 29-OCT-1999; 990S-0162142.  
  
Query Match 50.0%; Score 5; DB 21; Length 179;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 PVAPT 5  
Db 15 pvapt 19  
  
RESULT 39  
AAC59942  
ID AAC59942 standard; Protein; 179 AA.  
AC  
XX AAC59942;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEO ID NO: 77589.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX EPI033405-A2.  
PN  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 990S-0121825.  
PR 05-MAR-1999; 990S-0123180.  
PR 09-MAR-1999; 990S-0123548.  
PR 23-MAR-1999; 990S-0125788.  
PR 25-MAR-1999; 990S-0126264.  
PR 29-MAR-1999; 990S-0126785.  
PR 01-APR-1999; 990S-0127462.  
PR 06-APR-1999; 990S-0128234.  
PR 08-APR-1999; 990S-0128714.  
PR 16-APR-1999; 990S-0129845.  
PR 19-APR-1999; 990S-0130077.  
PR 21-APR-1999; 990S-0130449.  
PR 23-APR-1999; 990S-0130510.  
PR 28-APR-1999; 990S-0131449.  
PR 30-APR-1999; 990S-0132048.  
PR 30-APR-1999; 990S-0132407.  
PR 04-MAY-1999; 990S-0132484.  
PR 05-MAY-1999; 990S-0132485.  
PR 06-MAY-1999; 990S-0132486.  
PR 06-MAY-1999; 990S-0132487.  
PR 07-MAY-1999; 990S-0132683.  
PR 11-MAY-1999; 990S-0134256.  
PR 14-MAY-1999; 990S-0134218.  
PR 14-MAY-1999; 990S-0134219.  
PR 14-MAY-1999; 990S-0134221.  
PR 14-MAY-1999; 990S-0134370.  
PR 18-MAY-1999; 990S-0134768.  
PR 19-MAY-1999; 990S-0134941.  
PR 20-MAY-1999; 990S-0135124.

PR 21-MAY-1999; 990S-0135353.  
PR 24-MAY-1999; 990S-0135629.  
PR 25-MAY-1999; 990S-0136021.  
PR 27-MAY-1999; 990S-0136392.  
PR 28-MAY-1999; 990S-0136782.  
PR 01-JUN-1999; 990S-0137222.  
PR 03-JUN-1999; 990S-0137528.  
PR 04-JUN-1999; 990S-0137502.  
PR 07-JUN-1999; 990S-0137724.  
PR 08-JUN-1999; 990S-0138094.  
PR 10-JUN-1999; 990S-0138540.  
PR 10-JUN-1999; 990S-0138847.  
PR 14-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139452.  
PR 16-JUN-1999; 990S-0139453.  
PR 17-JUN-1999; 990S-0139492.  
PR 18-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139455.  
PR 18-JUN-1999; 990S-0139456.  
PR 18-JUN-1999; 990S-0139457.  
PR 18-JUN-1999; 990S-0139458.  
PR 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 18-JUN-1999; 990S-0139763.  
PR 21-JUN-1999; 990S-0139817.  
PR 22-JUN-1999; 990S-0139899.  
PR 23-JUN-1999; 990S-0140353.  
PR 23-JUN-1999; 990S-0140354.  
PR 24-JUN-1999; 990S-0140695.  
PR 28-JUN-1999; 990S-0140823.  
PR 29-JUN-1999; 990S-0140991.  
PR 30-JUN-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141842.  
PR 02-JUL-1999; 990S-0142154.  
PR 06-JUL-1999; 990S-0142055.  
PR 06-JUL-1999; 990S-0142390.  
PR 08-JUL-1999; 990S-0142803.  
PR 09-JUL-1999; 990S-0142920.  
PR 12-JUL-1999; 990S-0142977.  
PR 13-JUL-1999; 990S-0143542.  
PR 14-JUL-1999; 990S-0143624.  
PR 15-JUL-1999; 990S-0144005.  
PR 16-JUL-1999; 990S-0144085.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 19-JUL-1999; 990S-0144335.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144632.  
PR 21-JUL-1999; 990S-0144884.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 21-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145085.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145089.  
PR 22-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 23-JUL-1999; 990S-0145224.  
PR 26-JUL-1999; 990S-0145276.  
PR 27-JUL-1999; 990S-0145913.  
PR 27-JUL-1999; 990S-0145918.  
PR 27-JUL-1999; 990S-0145919.  
PR 28-JUL-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.

PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134376.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141847.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143547.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144637.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151087.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155133.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0156659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160811.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.



```

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159320.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161892.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match          50.0%; Score 5; DB 21; Length 152;
Best Local Similarity 100.0%; Pred.No.2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PVAPT 5
    |||||
Db 95 pvapt 99

```

```

RESULT 37
AAW89904
ID AAW89904 standard; Protein; 174 AA.
AC
XX AAW89904;
AC
XX
DT 18-FEB-1999 (first entry)
XX
DE Antigen from cluster 31.
XX
KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
XX
OS Helicobacter pylori.
XX
PN W09849314-A2.
XX
PD 05-NOV-1998.
XX
PF 27-APR-1998; 98MO-US08487.
XX
PR 14-OCT-1997; 97US-0061958.
PR 25-APR-1997; 97US-0045107.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Chow TP, FLY KE, LIM MY, Mcatee CP;
XX
DR WPI; 1999-009433/01.
XX
PT New Helicobacter pylori antigens and related nucleic acid sequences
- useful in serological diagnosis and protective vaccines, providing

```

```

PT Long-lasting immune response
XX
PS Claim 1; Page 240; 402pp; English.
XX
CC The present sequence represents a Helicobacter pylori antigenic protein
CC that is characterised by immunoreactivity with H. pylori positive
CC antisera. The proteins are highly immunogenic and induce a long-lasting
CC immune response that persists even after antimicrobial treatment. In
CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are
CC highly sensitive and specific. The specification also describes 69
CC previously unrecognised immunogenic cluster families. H. pylori antigens
CC are used to detect H. pylori-specific antibodies, for diagnosing
CC infection or to confirm eradication of infection, and in vaccines to
CC protect against H. pylori infection and related diseases (gastritis,
CC peptic ulcer, gastric adenocarcinoma/lymphoma).
XX
SQ Sequence 174 AA;

```

```

Query Match          50.0%; Score 5; DB 20; Length 174;
Best Local Similarity 100.0%; Pred.No.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 QEVKK 10
    |||||
Db 8 qevkk 12

```

```

RESULT 38
AAG55986
ID AAG55986 standard; Protein; 179 AA.
XX
AC AAG55986;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 71883.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0133285.
PR 06-MAY-1999; 99US-0133286.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.

```

PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141847.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147712.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149471.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149907.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.

PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142877.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145318.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150684.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.

PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 5; DB 21; Length 148;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APR0E 7  
|||||  
Db 93 aptge 97

RESULT 36  
AAG54453  
ID AAG54453 standard; protein; 152 AA.

XX AAG54453;

XX 18-OCT-2000 (first entry)

XX Zea mays protein fragment seq ID NO: 69433.

XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence; corn.

XX Zea mays subsp. mays.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123358.

PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158366.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 5; DB 21; Length 148;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 APTOE 7  
Db 93 aptge 97

RESULT 35  
AAG53842  
ID AAG53842 standard; Protein. 148 AA.

XX AAG53842;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 68587.

KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX

PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.



```
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141844.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145066.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145228.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145227.
PR 27-JUL-1999; 99US-0145216.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149724.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 5; DB 21; Length 142;
Best Local Similarity 100.0%; Pred. No. 2,5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 APTQE 7
Db 87 aptqe 91

RESULT 34
AAG08988
ID AAG08988 standard; Protein; 148 AA.
AC AAG08988;
XX
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6742.
XX
KW Protein Identification; signal transduction pathway; metabolic pathway;
```

PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149920.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 18-OCT-1999; 99US-0159638.  
PR 21-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 5; DB 21; Length 142;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APR0E 7  
Db 87 aptqe 91

RESULT 33  
AAC53843  
ID AAC53843 standard; Protein: 142 AA.  
XX

AC AAG53843;  
XX 18-OCT-2000 (first entry)  
DT  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 66588.  
DE  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 66588.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135553.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137724.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138647.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.

Db 42 gevkk 46

RESULT 32

ACG08989

ACG08989 standard; Protein: 142 AA.

XX

AC ACG08989;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 6743.

XX

KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EPI033405-A2.

XX

PD 06-SEP-2000.

XX

PE 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131444.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139452.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 24-JUN-1999; 99US-0140354.

PR 28-JUN-1999; 99US-0140695.

PR 29-JUN-1999; 99US-0140823.

PR 30-JUN-1999; 99US-0140991.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144864.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 10-AUG-1999; 99US-0147935.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148365.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.



```

PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147040.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147303.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149829.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150584.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151308.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0156559.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.

```

```

PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

```

```

Query Match          50.0%; Score 5; DB 21; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 6 QEVKK 10
Db 63 qevkk 67

```

## RESULT 31

```

ID AAY36760 standard; Protein; 135 AA.

```

```

AC AAY36760;

```

```

DT 07-OCT-1999 (first entry)

```

```

DE Amino acid sequence of a Chlamydia trachomatis protein.

```

```

KW Vaccine; eye disease; conventional trachoma; nongonococcal urethritis; epididymitis; genital disease; peritphalitis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.

```

```

XX Chlamydia trachomatis.

```

```

PN W09928475-A2.

```

```

PD 10-JUN-1999.

```

```

PF 27-NOV-1998; 98WO-IB01939.

```

```

PR 04-NOV-1998; 98US-0107077.
PR 28-NOV-1997; 97FR-0015041.
PR 17-DEC-1997; 97FR-0016034.

```

```

XX (GEST ) GENSET.

```

```

PA Griffais R;

```

```

PT WPI; 1999-371125/31.

```

```

XX Genome sequence of Chlamydia trachomatis

```

```

XX Disclosure; Page 662-663; 1755pp; English.

```

```

XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nongonococcal urethritis,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC peritphalitis, Bartholinitis; pneumopathy in breast feeding infants;
CC and venereal lymphogranulomatosis. The polypeptides of the invention
CC may be of use in treating these diseases.
XX
SQ Sequence 135 AA;

```

```

Query Match          50.0%; Score 5; DB 20; Length 135;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 6 QEVKK 10
Db 63 qevkk 67

```

C. pneumoniae causes respiratory disease such as pneumonia and  
CC bronchitis and is thought to be a contributing factor in heart  
CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
CC frames of the C. pneumoniae genome (see AAI34564-Y35879) can be used in  
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
CC nucleotide sequences can also be used as immunogenic compositions,  
CC especially where the vector directs the expression of a neutralising  
CC epitope of C. pneumoniae.

SQ Sequence 132 AA:

Query Match Best Local Similarity 50.0%; Score 5; DB 20; Length 132;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 QEVKK 10  
|||||  
Db 39 gevkk 43

RESULT 30  
AAG59953  
ID AAG59953 standard; Protein: 134 AA.

AC AAG59953;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 77605.

KM Protein identification: signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 29-MAR-1999; 99US-0126264.

PR 01-APR-1999; 99US-0126785.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140699.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141844.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0145086.  
PR 19-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.

PR 14-OCT-1999; 9905-0159329.  
 PR 14-OCT-1999; 9905-0159330.  
 PR 14-OCT-1999; 9905-0159331.  
 PR 14-OCT-1999; 9905-0159637.  
 PR 14-OCT-1999; 9905-0159638.  
 PR 14-OCT-1999; 9905-0159584.  
 PR 21-OCT-1999; 9905-0160741.  
 PR 21-OCT-1999; 9905-0160767.  
 PR 21-OCT-1999; 9905-0160770.  
 PR 21-OCT-1999; 9905-0160776.  
 PR 21-OCT-1999; 9905-0160814.  
 PR 21-OCT-1999; 9905-0160815.  
 PR 22-OCT-1999; 9905-0160980.  
 PR 22-OCT-1999; 9905-0160981.  
 PR 22-OCT-1999; 9905-0160989.  
 PR 25-OCT-1999; 9905-0161404.  
 PR 25-OCT-1999; 9905-0161405.  
 PR 25-OCT-1999; 9905-0161406.  
 PR 26-OCT-1999; 9905-0161359.  
 PR 26-OCT-1999; 9905-0161360.  
 PR 26-OCT-1999; 9905-0161361.  
 PR 28-OCT-1999; 9905-0161920.  
 PR 28-OCT-1999; 9905-0161992.  
 PR 28-OCT-1999; 9905-0161993.  
 PR 29-OCT-1999; 9905-0162142.

Query Match 50.0%; Score 5; DB 21; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PVAPT 5  
 |||||

Db 68 pvapt 72

# RESULT 28

AAW77657  
 ID AAW77657 standard; Protein: 130 AA.

XX AC AAW77657;

XX DT 30-OCT-1998 (first entry)

XX DE Staphylococcus aureus protein of unknown function.

XX KM Staphylococcus aureus protein: immune response induction; eye infection;

KM antibody production; T-cell immune response; gastrointestinal infection;

KM respiratory infection; inhibitory; bacterial infection; cardiac infection;

KM central nervous system; kidney infection; urinary tract infection;

KM antimicrobial compound identification; broad spectrum antibiotic;

XX OS Staphylococcus aureus.

XX FH Key Location/Qualifiers

XX FT MISC-difference 1..130

XX FT "residues designated X are unspecified, and

XX represented as Xaa in the specification"

XX PN EP841394-A2.

XX PD 13-MAY-1998.

XX PF 24-SEP-1997; 97EP-0307485.

XX PR 24-SEP-1996; 96US-0027032.

XX PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX PI Black MT, Burnham MKR, Hodgson JE, Knowles DJC;

PI Lonetto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M;

PI Ward JM;

XX XH  
 DR WPI: 1998-252940/23.  
 DR N-PSDB; AAV53451.  
 XX  
 PT New nucleic acid sequences from Staphylococcus aureus MCH029 -  
 PT useful in vaccines and for treatment of bacterial infections of e.g.  
 PT respiratory tract and central nervous system  
 XX  
 PS Claim 11; Page 312; 390pp; English.

CC This sequence represents a Staphylococcus aureus protein of unknown  
 CC function, and is encoded by a DNA sequence of the invention.  
 CC The DNA sequences were isolated from Staphylococcus aureus MCH029  
 CC (NCIMB 40771). Host cells containing the DNA sequences are used to  
 CC produce polypeptides or fragments. The proteins are used in the treatment  
 CC of disease, for inducing an immune response by administering them, to  
 CC produce antibody and/or T-cell immune response. Antagonists of the  
 CC proteins are used for the inhibition of bacterial polypeptides.  
 CC Conditions which may be treated include bacterial infections, especially  
 CC respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,  
 CC urinary tract, skin, bones and joints. The proteins can also be used to  
 CC identify antimicrobial compounds which are broad spectrum antibiotics,  
 CC especially useful in the treatment of H. pylori infection.

XX SQ Sequence 130 AA;

Query Match 50.0%; Score 5; DB 19; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 VAPRO 6  
 |||||

Db 108 vaptq 112

# RESULT 29

AAV35155  
 ID AAV35155 standard; Protein: 132 AA.

XX AC AAV35155;

XX DT 13-SEP-1999 (first entry)

XX DE Chlamydia pneumoniae transmembrane protein sequence.

XX KM Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;

KM sinusitis; purulent otitis media; erythema nodosum; pharyngitis;

XX OS Chlamydia pneumoniae.

XX PN W09927105-A2.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-IR01890.

XX PR 04-NOV-1998; 98US-0107078.

XX PR 21-NOV-1997; 97FR-0014673.

XX PA (GEST ) GENSET.

XX PI Giffels R;

XX DR WPI: 1999-357842/30.

XX PT Genome sequence of Chlamydia pneumoniae

XX PS Page 1021; Disclosure; 1912pp; English.

XX CC AAV34584-Y35879 represent the proteins encoded by all the open reading

CC frames in the complete genome (see AAV31990) of Chlamydia pneumoniae.

PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130044.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142153.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151067.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154038.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.

```

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147360.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

```

```

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match          50.0%; Score 5; DB 21; Length 121;
Best Local Similarity 100.0%; Pred. No. 2; 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 PVAPT 5
      |||||
Db      64 pvapt 68

```

```

RESULT 27
AAC54454
ID      AAG54454 standard; Protein; 125 AA.

```

```

AC      AAG54454;

```

```

XX      18-OCT-2000 (first entry)

```

```

DT      Zea mays protein fragment SEQ ID NO: 69434.

```

```

XX      Protein identification: signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KM      termination sequence; corn.

```

```

XX      Zea mays subsp. mays.

```

```

XX      EPI033405-A2.

```

```

XX      06-SEP-2000.

```

```

PF      25-FEB-2000; 2000EP-0301439.

```

```

XX      25-FEB-1999; 99US-0121825.

```

```

PR      05-MAR-1999; 99US-0123180.

```

```

PR      09-MAR-1999; 99US-0123548.

```

```

PR      23-MAR-1999; 99US-0125788.

```

```

PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128234.
PR      08-APR-1999; 99US-0128714.

```

PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158368.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 5; DB 21; Length 114;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 OEYKK 10  
| | | | |  
Db 43 gevkk 47

## RESULT 26

AAG54455  
ID AAG54455 standard; Protein: 121 AA.

AC AAG54455;

DT 18-OCT-2000 (first entry)

DE Zea mays protein fragment SEQ ID NO: 69435.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.

OS Zea mays subsp. mays.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130044.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132047.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141844.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.

XX 06-SEP-2000.  
PD  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 990S-0121825.  
PR 05-MAR-1999; 990S-0123180.  
PR 09-MAR-1999; 990S-0123548.  
PR 23-MAR-1999; 990S-0125788.  
PR 25-MAR-1999; 990S-0126264.  
PR 29-MAR-1999; 990S-0126785.  
PR 01-APR-1999; 990S-0127462.  
PR 06-APR-1999; 990S-0128234.  
PR 08-APR-1999; 990S-0128714.  
PR 16-APR-1999; 990S-0129845.  
PR 19-APR-1999; 990S-0130077.  
PR 21-APR-1999; 990S-0130049.  
PR 23-APR-1999; 990S-0130510.  
PR 28-APR-1999; 990S-0130891.  
PR 30-APR-1999; 990S-0132048.  
PR 30-APR-1999; 990S-0132407.  
PR 04-MAY-1999; 990S-0132484.  
PR 05-MAY-1999; 990S-0132485.  
PR 06-MAY-1999; 990S-0132486.  
PR 06-MAY-1999; 990S-0132487.  
PR 07-MAY-1999; 990S-0132863.  
PR 11-MAY-1999; 990S-0134256.  
PR 14-MAY-1999; 990S-0134218.  
PR 14-MAY-1999; 990S-0134219.  
PR 14-MAY-1999; 990S-0134221.  
PR 14-MAY-1999; 990S-0134370.  
PR 18-MAY-1999; 990S-0134768.  
PR 19-MAY-1999; 990S-0134941.  
PR 20-MAY-1999; 990S-0135124.  
PR 21-MAY-1999; 990S-0135353.  
PR 24-MAY-1999; 990S-0135629.  
PR 25-MAY-1999; 990S-0136021.  
PR 27-MAY-1999; 990S-0136392.  
PR 28-MAY-1999; 990S-0136782.  
PR 01-JUN-1999; 990S-0137222.  
PR 03-JUN-1999; 990S-0137528.  
PR 04-JUN-1999; 990S-0137502.  
PR 07-JUN-1999; 990S-0137724.  
PR 08-JUN-1999; 990S-0138094.  
PR 10-JUN-1999; 990S-0138540.  
PR 10-JUN-1999; 990S-0138847.  
PR 14-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139452.  
PR 16-JUN-1999; 990S-0139453.  
PR 17-JUN-1999; 990S-0139492.  
PR 18-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139455.  
PR 18-JUN-1999; 990S-0139456.  
PR 18-JUN-1999; 990S-0139457.  
PR 18-JUN-1999; 990S-0139458.  
PR 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 18-JUN-1999; 990S-0139763.  
PR 21-JUN-1999; 990S-0139817.  
PR 22-JUN-1999; 990S-0139899.  
PR 23-JUN-1999; 990S-0140353.  
PR 23-JUN-1999; 990S-0140354.  
PR 24-JUN-1999; 990S-0140695.  
PR 28-JUN-1999; 990S-0140823.  
PR 29-JUN-1999; 990S-0140991.  
PR 30-JUN-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141842.  
PR 01-JUL-1999; 990S-0142154.  
  
PR 02-JUL-1999; 990S-0142055.  
PR 06-JUL-1999; 990S-0142390.  
PR 08-JUL-1999; 990S-0142803.  
PR 09-JUL-1999; 990S-0142820.  
PR 12-JUL-1999; 990S-0142877.  
PR 13-JUL-1999; 990S-0143542.  
PR 14-JUL-1999; 990S-0143624.  
PR 15-JUL-1999; 990S-0144005.  
PR 16-JUL-1999; 990S-0144085.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 19-JUL-1999; 990S-0144335.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144632.  
PR 20-JUL-1999; 990S-0144684.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 21-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145085.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145089.  
PR 22-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 23-JUL-1999; 990S-0145324.  
PR 26-JUL-1999; 990S-0145376.  
PR 27-JUL-1999; 990S-0145913.  
PR 27-JUL-1999; 990S-0145918.  
PR 27-JUL-1999; 990S-0145919.  
PR 28-JUL-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.  
PR 02-AUG-1999; 990S-0146388.  
PR 02-AUG-1999; 990S-0146389.  
PR 03-AUG-1999; 990S-0147038.  
PR 04-AUG-1999; 990S-0147204.  
PR 04-AUG-1999; 990S-0147302.  
PR 05-AUG-1999; 990S-0147192.  
PR 05-AUG-1999; 990S-0147260.  
PR 06-AUG-1999; 990S-0147303.  
PR 06-AUG-1999; 990S-0147416.  
PR 09-AUG-1999; 990S-0147493.  
PR 09-AUG-1999; 990S-0147935.  
PR 10-AUG-1999; 990S-0148171.  
PR 11-AUG-1999; 990S-0148319.  
PR 12-AUG-1999; 990S-0148341.  
PR 13-AUG-1999; 990S-0148565.  
PR 13-AUG-1999; 990S-0148684.  
PR 16-AUG-1999; 990S-0149368.  
PR 17-AUG-1999; 990S-0149375.  
PR 18-AUG-1999; 990S-0149426.  
PR 20-AUG-1999; 990S-0149722.  
PR 20-AUG-1999; 990S-0149723.  
PR 20-AUG-1999; 990S-0149929.  
PR 23-AUG-1999; 990S-0149902.  
PR 23-AUG-1999; 990S-0149930.  
PR 25-AUG-1999; 990S-0150566.  
PR 26-AUG-1999; 990S-0150884.  
PR 27-AUG-1999; 990S-0151065.  
PR 27-AUG-1999; 990S-0151066.  
PR 27-AUG-1999; 990S-0151080.  
PR 30-AUG-1999; 990S-0151303.  
PR 31-AUG-1999; 990S-0151348.  
PR 01-SEP-1999; 990S-0151330.  
PR 07-SEP-1999; 990S-0152363.  
PR 10-SEP-1999; 990S-0153070.  
PR 13-SEP-1999; 990S-0153758.  
PR 15-SEP-1999; 990S-0154018.  
PR 16-SEP-1999; 990S-0154039.  
PR 20-SEP-1999; 990S-0154779.

```
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 5; DB 21; Length 114;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 OEKK 10
Db 43 qevkk 47

RESULT 25
AA65954
ID AA65954 standard; Protein; 114 AA.
XX
XX AA65954;
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 77606.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
```



PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151303.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 5; DB 21; Length 110;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 QEWKK 10  
Db 39 qevkk 43

RESULT 24  
AAG12210  
ID AAG12210 standard; Protein; 114 AA.  
XX  
AC AAG12210;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 11234.  
XX

KW Protein identification: signal transduction pathway; metabolic pathway;  
KW hybridisation assay: genetic mapping; gene expression control; promoter;  
KW termination sequence.  
OS Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
XX  
XX PD 06-SEP-2000.  
XX  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 28-APR-1999; 99US-0130891.  
XX 30-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 06-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134321.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
XX 28-MAY-1999; 99US-0136782.  
XX 01-JUN-1999; 99US-0137522.  
XX 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137502.  
XX 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138847.  
XX 14-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
XX 16-JUN-1999; 99US-0139453.  
XX 17-JUN-1999; 99US-0139492.  
XX 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.  
XX 18-JUN-1999; 99US-0139456.  
XX 18-JUN-1999; 99US-0139457.  
XX 18-JUN-1999; 99US-0139458.  
XX 18-JUN-1999; 99US-0139459.  
XX 18-JUN-1999; 99US-0139460.  
XX 18-JUN-1999; 99US-0139461.  
XX 18-JUN-1999; 99US-0139462.  
XX 18-JUN-1999; 99US-0139463.  
XX 18-JUN-1999; 99US-0139750.  
XX 18-JUN-1999; 99US-0139763.  
XX 21-JUN-1999; 99US-0139817.  
XX 22-JUN-1999; 99US-0139899.  
XX 23-JUN-1999; 99US-0140353.

XX AAG59955;  
AC 18-OCT-2000 (first entry)  
DT  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 77607.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hydrolisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX 25-FEB-1999; 99US-0121825.  
PR 09-MAR-1999; 99US-0123180.  
PR 23-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125786.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 11-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.

PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150984.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155186.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 50.0%; Score 5; DB 21; Length 110;  
Best Local Similarity 100.0%; Pred. No. 26+02; Indels 0; Gaps 0;  
Matches 5; Conservative 0; Mismatches 0;

Oy 6 OEYKK 10  
|||||  
Db 39 gevkk 43

RESULT 23  
AAG59955  
ID AAG59955 standard; Protein; 110 AA.

RESULT 21  
AAW75784  
ID AAW75784 standard; Protein: 107 AA.  
XX  
AC AAW75784;  
XX  
DT 18-JAN-1999 (first entry)  
XX  
DE Human lymphocyte surface receptor cytoplasmic domain.  
XX  
TACI; transmembrane activator and CAML-interactor;  
XX calcium signal-modulating cyclophilin ligand; human;  
XX lymphocyte surface receptor; human; B-cell; B lymphocyte;  
XX infection; cancer; rheumatoid arthritis; autoimmune disease;  
XX glomerulonephritis; immunosuppressive; graft versus host disease;  
XX transplant rejection; therapy; signal transduction.  
XX  
OS Homo sapiens.  
XX  
PN WO9839361-A1.  
XX  
PD 11-SEP-1998.  
XX  
PF 03-MAR-1998; 98WO-US04270.  
XX  
PR 03-MAR-1997; 97US-0810572.  
XX  
PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
XX  
PI Bram RJ, Von Bulow G;  
XX  
DR WPI: 1998-506346/43.  
XX N-PSDB; AAV57329.  
XX  
PT New isolated transmembrane activator protein - used to develop  
PT products for treating e.g. infections, cancers, autoimmune and  
PT inflammatory conditions, transplant rejection or graft-versus-host  
PT disease  
XX  
PS Claim 6; Page 72; 89p; English.  
XX  
XX This polypeptide comprises the C-terminal, i.e. cytoplasmic,  
XX domain of novel human transmembrane activator and CAML-interactor  
XX protein TACI (see AAW75783). TACI is a lymphocyte receptor protein  
XX that is involved in the calcium activation pathway. It is normally  
XX present in B-lymphocytes, and to a much lesser extent in immature  
XX T-lymphocytes, and can therefore be targeted to specifically  
XX regulate B cell responses without affecting T cell activity. The  
XX cytoplasmic domain of TACI mediates signal transduction via  
XX Ca2+-dependent and Ca2+-independent mechanisms and includes the  
XX CAML-binding region of TACI. Methods are claimed for identifying a  
XX ligand for TACI and for identifying immunosuppressive drugs that  
XX selectively block the action of B lymphocytes without affecting  
XX mature T lymphocytes. TACI can be activated to increase immune  
XX system activity, e.g. for treating infections or cancers. It can  
XX be blocked to provide immunosuppression, e.g. for treating  
XX autoimmune and inflammatory conditions such as immune complex-  
XX induced vasculitis, glomerulonephritis, haemolytic anaemia,  
XX myasthenia gravis, type II collagen-induced arthritis, experimental  
XX allergic and hyperacute xenograft rejection, rheumatoid arthritis,  
XX systemic lupus erythematosus, transplant rejection, cancer or  
XX graft versus host disease.  
XX  
SO Sequence 107 AA;

Query Match 50.0%; Score 5; DB 19; Length 107;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 APTOE 7

Db 53 aptge 57  
IIII  
RESULT 22  
AAG12211  
ID AAG12211 standard; Protein: 110 AA.  
XX  
AC AAG12211;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 11235.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123348.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 23-APR-1999; 99US-0130891.  
XX 28-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 30-APR-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 06-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
XX 28-MAY-1999; 99US-0136782.  
XX 01-JUN-1999; 99US-0137222.  
XX 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137502.  
XX 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138847.  
XX 14-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
XX 16-JUN-1999; 99US-0139453.  
XX 17-JUN-1999; 99US-0139492.  
XX 18-JUN-1999; 99US-0139454.

CC which are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Also, pathological  
CC conditions can be diagnosed by determining the amount of the new  
CC polypeptides in a sample or by determining the presence of mutations in  
CC the new polynucleotides. Specific uses are described for each of the 125  
CC polynucleotides, based on which tissues they are most highly expressed in  
CC (see AAX84933 for described uses).

XX  
SQ Sequence 87 AA;

Query Match

Best Local Similarity 50.0%; Score 5; DB 20; Length 87;  
Matches 5; Conservative 100.0%; Pred. No. 1.7e+02; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
| | | | |  
Db 60 pvapt 64

RESULT 19

AA82898  
ID AAY82898 standard; Protein: 101 AA.

AC AAY82898;

DT 24-JUL-2000 (first entry)

DE CUB domain from BMP-1 protein of mouse.

XX BPC-1; oncogene; oncogenic; cancer; prostate; bladder; antibody;  
KM antisenesc; vaccine; detection; prognosis; drug screening; BMP-1;  
XX mouse.

OS Mus musculus.

PN WO200009691-A2.

PD 24-FEB-2000.

PF 10-AUG-1999; 99WO-US18250.

PR 10-AUG-1998; 98US-0095982.

XX (UROG-) UROGENESYS INC.

PA (AFAR) AFAR D E.

PA (HUBE) HUBERT R S.

PA (LEON/) LEONG K.

PA (RAIT/) RAITANO A B.

PA (SAFE/) SAFRAN D C.

PA (JAKO/) JAKOBOVITS A.

PI Afar DE, Hubert RS, Leong K, Raitano AB, Safran DC, Jakobovits A;

XX WPI; 2000-206006/18.

XX New isolated BPC-1 polypeptides, useful for developing products for the

PT diagnosis, staging, prognosis and treatment of cancers, particularly

PT prostate or bladder cancer

XX Disclosure; Figure 3; 79pp; English.

XX BPC-1 polypeptides and polynucleotides can be used for the detection  
CC of BPC-1 polypeptides and polynucleotides in biological samples, this  
CC is particularly useful for detecting cancers expressing BPC-1, e.g.  
CC prostate cancer or bladder cancer. Antibodies directed against BPC-1  
CC or antisenesc polynucleotides can also be used for treating such cancers.  
CC The BPC-1 polypeptides can also be used in vaccines for treating or  
CC inhibiting the development of a cancer expressing BPC-1. The  
CC polypeptides and polynucleotides can also be used for detection,  
CC prognosis, drug screening and predicting susceptibility to developing  
CC cancer. In normal human tissues BPC-1 is only expressed in certain  
CC tissues of the brain, however, it is expressed at high levels in

CC prostate cancer cells and bladder cancer cells. The BPC-1 polypeptide  
CC comprises a CUB domain which is expressed in prostate and bladder  
CC carcinoma cells and which shows sequence similarity with CUB domains  
CC from other known proteins. This is a CUB domain from the BMP-1  
CC protein of a mouse. It shows 36.3% identity with the  
CC BPC-1 CUB domain 3 in 102 residues overlap.

XX  
SQ Sequence 101 AA;

Query Match

Best Local Similarity 50.0%; Score 5; DB 21; Length 101;  
Matches 5; Conservative 100.0%; Pred. No. 1.9e+02; Mismatches 0; Indels 0; Gaps 0;

OY 2 VAPQ 6  
| | | | |  
Db 23 vapq 27

RESULT 20

AAB72875  
ID AAB72875 standard; Protein: 102 AA.

AC AAB72875;

DT 09-MAY-2001 (first entry)

DE Human p53 A76T/V122A mutated peptide.

XX Human; p53 mutation; mutant; mutein; gene therapy; supertransactivating;  
KM cancer.

XX Homo sapiens.

OS Synthetic.

PN WO200109325-A2.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000WO-US20538.

PR 30-JUL-1999; 99US-0146634.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PA Resnick MA, Inga A;

PI WPI; 2001-123321/13.

XX The present invention provides a number of peptides derived from the

XX human p53 protein and containing specific mutations. These are useful for

CC isolating supertransactivating and toxic p53 mutants, and for inhibiting

CC cell growth and treating cancer. The present sequence is a peptide

CC derived from the wild-type human p53 protein shown in AAB72878.

CC Note: The present sequence is not shown in the specification but is

CC derived from that given in Genbank acc. no. X02469 (see AAB72878).

XX  
SQ Sequence 102 AA;

Query Match  
Best Local Similarity 50.0%; Score 5; DB 22; Length 102;  
Matches 5; Conservative 100.0%; Pred. No. 1.9e+02; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
| | | | |  
Db 23 pvapt 27

OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH Misc-difference 27  
 FT /label= unknown  
 FT  
 XX MO9856804-A1.  
 XX  
 XX 17-DEC-1998.  
 XX  
 XX 11-JUN-1998; 98WO-US12125.  
 XX  
 XX 02-OCT-1997; 97US-0061060.  
 XX 13-JUN-1997; 97US-0049547.  
 XX 13-JUN-1997; 97US-0049548.  
 XX 13-JUN-1997; 97US-0049549.  
 XX 13-JUN-1997; 97US-0049550.  
 XX 13-JUN-1997; 97US-0049606.  
 XX 13-JUN-1997; 97US-0049607.  
 XX 13-JUN-1997; 97US-0049608.  
 XX 13-JUN-1997; 97US-0049609.  
 XX 13-JUN-1997; 97US-0049610.  
 XX 13-JUN-1997; 97US-0049611.  
 XX 13-JUN-1997; 97US-0050566.  
 XX 13-JUN-1997; 97US-0050901.  
 XX 13-JUN-1997; 97US-0052989.  
 XX 08-JUL-1997; 97US-0051919.  
 XX 18-AUG-1997; 97US-0055984.  
 XX 12-SEP-1997; 97US-0058665.  
 XX 12-SEP-1997; 97US-0058666.  
 XX 12-SEP-1997; 97US-0058669.  
 XX 12-SEP-1997; 97US-0058750.  
 XX 12-SEP-1997; 97US-0058971.  
 XX 12-SEP-1997; 97US-0058972.  
 XX 12-SEP-1997; 97US-0058975.  
 XX 02-OCT-1997; 97US-0060834.  
 XX 02-OCT-1997; 97US-0060841.  
 XX 02-OCT-1997; 97US-0060844.  
 XX 02-OCT-1997; 97US-0060865.  
 XX 02-OCT-1997; 97US-0061059.  
 XX  
 XX (HUMA-) HUMAN GENOME SCT INC.  
 XX  
 XX Brewer LA, Ebner R, Ferris AM, Feng P, Greene JM, Lafleur DW;  
 XX Moore PA, Nl J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;  
 XX Yu GL;  
 XX WPI: 1999-080881/07.  
 XX N-PSDB: AAX04315.  
 XX  
 XX New isolated human genes and the secreted polypeptides they encode -  
 XX useful for diagnosis and treatment of e.g. cancers, neurological  
 XX disorders, immune diseases, inflammation or blood disorders  
 XX  
 XX Claim 11: Page 266; 380pp; English.  
 XX  
 XX This sequence represents a secreted human protein encoded by the gene  
 XX clone detailed in the descriptor line. The gene can be used to generate  
 XX fusion proteins by linking to the gene to a human immunoglobulin Fc  
 XX portion (e.g. AAX04302) for increasing the stability of the fused  
 XX protein as compared to the human protein only.  
 XX The invention relates to 86 novel genes and their fragments (nucleic  
 XX acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)  
 XX which are useful for preventing, treating or ameliorating medical  
 XX conditions e.g. by protein or gene therapy. Also, pathological  
 XX conditions can be diagnosed by determining the amount of the new  
 XX polypeptides in a sample or by determining the presence of mutations in  
 XX the new polynucleotides. Specific uses are described for each of the 86  
 XX polynucleotides, based on which tissues they are most highly expressed in  
 XX (see AAX04311 for described uses).  
 XX  
 XX Sequence 27 AA:

Query Match 50.0%; Score 5; DB 20; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 5; \*Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 PVAPT 5.  
 Db 9 pvaPT 13  
 RESULT 18  
 AAY27569  
 ID AAY27569 standard; Protein; 87 AA.  
 XX  
 AC AAY27569;  
 XX  
 DT 30-JUL-1999 (first entry)  
 XX  
 DE Human secreted protein encoded by gene No. 3.  
 XX  
 XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
 XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
 XX developmental abnormality; foetal deficiency; blood; allergy; renal;  
 XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 XX inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
 XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
 XX  
 OS Homo sapiens.  
 XX  
 XX MO9924836-A1.  
 XX  
 XX 20-MAY-1999.  
 XX  
 XX 04-NOV-1998; 98WO-US23435.  
 XX  
 XX 17-NOV-1997; 97US-0066100.  
 XX 07-NOV-1997; 97US-0064900.  
 XX 07-NOV-1997; 97US-0064908.  
 XX 07-NOV-1997; 97US-0064911.  
 XX 07-NOV-1997; 97US-0064912.  
 XX 07-NOV-1997; 97US-0064983.  
 XX 07-NOV-1997; 97US-0064984.  
 XX 07-NOV-1997; 97US-0064985.  
 XX 07-NOV-1997; 97US-0064987.  
 XX 07-NOV-1997; 97US-0064988.  
 XX 17-NOV-1997; 97US-0066090.  
 XX 17-NOV-1997; 97US-0066094.  
 XX 17-NOV-1997; 97US-0066095.  
 XX 17-NOV-1997; 97US-0066089.  
 XX  
 XX (HUMA-) HUMAN GENOME SCT INC.  
 XX  
 XX Carter KC, Ebner R, Endress GA, Feng P, Janat F;  
 XX Kyaw H, Lafleur DW, Moore PA, Nl J, Olsen HS, Rosen CA;  
 XX Ruben SM, Shi Y, Soppet DR, Wei Y;  
 XX WPI: 1999-337740/28.  
 XX N-PSDB: AAX84935.  
 XX  
 XX New human secreted proteins and coding sequences useful for treating  
 XX disorders of the immune system and hyperproliferative disorders  
 XX  
 XX Claim 11: Page 350; 507pp; English.  
 XX  
 XX This sequence represents a secreted human protein encoded by the gene  
 XX clone detailed in the descriptor line. The gene can be used to generate  
 XX fusion proteins by linking to the gene to a human immunoglobulin Fc  
 XX portion (e.g. AAX84924) for increasing the stability of the fused  
 XX protein as compared to the human protein only.  
 XX The invention relates to 125 novel genes and their fragments (nucleic  
 XX acid sequences: AAX84933-X85057; amino acid sequences AAY27567-Y27933)

CC from a human immunodeficiency virus type 1 (HIV-1) protein. The  
 CC immunogenic 11peptide micelles are used in vaccines, especially  
 CC against HIV, hepatitis B virus (HBV), papilloma viruses, p53, melanoma  
 CC or Plasmodium falciparum malaria.

SO Sequence 9 AA;

Query Match 50.0%; Score 5; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TQEVK 9  
 |||||  
 Db 3 tqevk 7

RESULT 15

AAR54624  
 ID AAR54624 standard; Protein; 11 AA.

XX AAR54624;

XX 16-JUN-1994 (first entry)

DE Listeria innocua p60 peptide epitope.

XX Antibodies; immunoassay; conjugate.

OS Synthetic.

PN DE4318450-A.

PD 16-DEC-1993.

XX 03-JUN-1993; 93DE-4318450.

XX 11-JUN-1992; 92DE-421911.

PR 25-NOV-1992; 92DE-4239567.

XX (MERE ) MERCK PATENT GMBH.

PI Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;  
 PI Hofmann G, Hubert A, Goebel W, Koehler S;

XX WPI; 1993-406956/51.

PT New primers for PCR detection of Listeria - including individual  
 PT species, also new peptide(s) for raising antibodies for  
 PT immunochemical detection

XX PS Disclosure; Fig 5; 19pp; German.

CC The sequence is that of a Listeria innocua p60 peptide epitope which  
 CC which may be used in the prodn. of antibodies for the detection  
 CC of Listeria by immunoassay (partic. ELISA). It may be used as  
 CC part of a method that allows determination of individual Listeria  
 CC species.

SO Sequence 11 AA;

Query Match 50.0%; Score 5; DB 14; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 QEVKK 10  
 |||||  
 Db 7 qevkk 11

RESULT 16  
 AAB72848

ID AAB72848 standard; Protein; 11 AA.

XX AAB72848;

XX 09-MAY-2001 (first entry)

DE Human p53 A76r mutated peptide.

KW Human: p53 mutation; mutant; mutein; gene therapy; supertransactivating;  
 KW cancer.

OS Homo sapiens.

PN WO200109325-A2.

PD 08-FEB-2001.

XX 28-JUL-2000; 2000WO-US20538.

XX 30-JUL-1999; 99US-0146634.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Resnick MA, Inga A;

XX WPI; 2001-123321/13.

PT New isolated mutated human p53 polypeptides for inducing toxicity in a  
 PT cell, treating cancer and identifying compounds that mimic toxic or  
 PT supertransactivating mutations -

PS Claim 6; Page -; 14pp; English.

CC The present invention provides a number of peptides derived from the  
 CC human p53 protein and containing specific mutations. These are useful for  
 CC isolating supertransactivating and toxic p53 mutants, and for inhibiting  
 CC cell growth and treating cancer. The present sequence is a peptide  
 CC derived from the wild-type human p53 protein shown in AAB72878.  
 CC Note: The present sequence is not shown in the specification but is  
 CC derived from that given in Genbank acc. no. X02469 (see AAB72878).

SO Sequence 11 AA;

Query Match 50.0%; Score 5; DB 22; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPT 5  
 |||||  
 Db 3 pvapt 7

RESULT 17  
 AAM78130  
 ID AAM78130 standard; Protein; 27 AA.

XX AAM78130;

XX 13-APR-1999 (first entry)

DE Human secreted protein encoded by gene 5 clone HPBDD36.

KW Human: secreted protein; fusion protein; gene therapy; protein therapy;  
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KW inflammation; ischaemic shock; Alzheimer's disease; osteoclast; AIDS;  
 KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
 KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

PT separated by peptide spacers that impart hydrophilicity, useful in  
 PT vaccines -  
 PS Disclosure: Page 23; 35pp; French.  
 XX  
 CC The invention relates to the generation of a lipopeptide comprising at  
 CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)  
 CC epitope and at least one lipid residue with (i) the epitopes and lipid  
 CC portion and (ii) the epitopes, being separated independently by peptide  
 CC spacers. These spacers comprise sequences of amino acids which carry an  
 CC overall electrical charge in neutral media to ensure that the  
 CC lipopeptide is hydrophilic. The peptides AAY5301-Y53549 represents  
 CC examples of peptide epitopes used to generate the lipopeptides. These are  
 CC used in therapeutic or prophylactic compositions and vaccines to induce  
 CC specific immune responses against human immunodeficiency, hepatitis B or  
 CC papilloma viruses; p53 of melanoma or the malaria parasite.  
 XX  
 SQ Sequence 9 AA:  
 Query Match 50.0%; Score 5; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 TOEVR 9  
 |||||  
 Db 3 tqevk 7  
 RESULT 13  
 AAY40367  
 ID AAY40367 standard; peptide: 9 AA.  
 XX  
 AC AAY40367;  
 DT 19-NOV-1999 (first entry)  
 XX  
 DE Amino acid sequence of a HIV-1 epitope.  
 XX  
 KW Cytotoxic T cell; T lymphocyte; CD8+ epitope; T helper cell;  
 KW CD4+ epitope; B epitope; lipopeptide; interferon gamma; adjuvant;  
 KW vaccine; tumor; infection; immune response; cytokine profile;  
 KW acquired immune deficiency syndrome; papilloma; cancer; hepatitis;  
 KW autoimmune disease.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN FR2774687-A1.  
 PD 13-AUG-1999.  
 XX  
 PF 06-FEB-1998; 98FR-0001439.  
 XX  
 PR 06-FEB-1998; 98FR-0001439.  
 XX  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (INSP ) INST PASTEUR LILLE.  
 PI Thiam K, Guillet JG, Ver. Waerde C, Auriault C, Gras MH, Loding E;  
 PI WPI; 1999-510734/43.  
 DR  
 XX New lipopeptide comprising C-terminal interferon-gamma fragment with  
 PT attached lipophilic groups, used as interferon mimic, e.g. for treating  
 PT cancer or virus infection  
 XX  
 PS Disclosure: Page 41; 53pp; French.  
 XX  
 CC AAY40123-140379 represent epitopes that are able to activate cytotoxic  
 CC T lymphocytes (CD8+ epitopes), T helper cells (CD4+ epitopes), or  
 CC B lymphocytes recognized by corresponding antibodies. The epitopes may be  
 CC used in the composition of the invention. The specification describes a  
 CC lipopeptide that has a peptide part derived from mammalian interferon

CC gamma (IFN $\gamma$ ) and one or more lipophilic parts comprising a linear or  
 CC branched, (un)saturated 4-20C hydrocarbonyl chain or a steroid. The  
 CC lipopeptide mimics the activity of IFN $\gamma$ . Compositions comprising the  
 CC lipopeptide are used to treat or prevent any condition that responds  
 CC to IFN $\gamma$ , and as adjuvant for vaccines (particularly those directed  
 CC against tumors, viral or parasitic infections), to stimulate or  
 CC (re)orient the immune response between types 1 and 2 cytokine profiles.  
 CC Particular applications are treatment of infections (particularly  
 CC viral, e.g. acquired immune deficiency syndrome, papilloma (cancer) and  
 CC hepatitis, but also bacterial, fungal, parasitic or helminth); cancers  
 CC (particularly of kidney, cutaneous T cells or ovary, chronic  
 CC myelogenous leukemia or mesothelioma), allergy; and autoimmune  
 CC diseases.  
 XX  
 SQ Sequence 9 AA:  
 Query Match 50.0%; Score 5; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 TOEVR 9  
 |||||  
 Db 3 tqevk 7  
 RESULT 14  
 AAY26859  
 ID AAY26859 standard; peptide: 9 AA.  
 XX  
 AC AAY26859;  
 DT 14-SEP-1999 (first entry)  
 XX  
 DE HIV-derived lipopeptide epitope #129 for mixed micelles.  
 XX  
 KW Micelle; microaggregate; induction; immune response; lipopeptide; CTL;  
 KW cytotoxic T-lymphocyte; epitope; lipid; helper T-lymphocyte; HTL; HBV;  
 KW tetanus; toxin; vaccine; HIV; hepatitis B virus; papilloma virus; p53;  
 KW melanoma; Plasmodium falciparum; malaria.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN FR2771640-A1.  
 PD 04-JUN-1999.  
 XX  
 PF 03-DEC-1997; 97FR-0015246.  
 XX  
 PR 03-DEC-1997; 97FR-0015246.  
 XX  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (INSP ) INST PASTEUR LILLE.  
 PI Bossus M, Bourgault VI, Gras-Masse H, Guillet JG, Lippens G;  
 PI Tartar A, Wieruszski JM;  
 PI WPI; 1999-349509/30.  
 DR  
 XX Immunogenic lipopeptide micelles - comprising lipopeptides  
 PT containing cytotoxic and helper T-lymphocyte epitopes  
 PT  
 PS Disclosure: Page 36; 60pp; French.  
 XX  
 CC The invention relates to the generation of mixed micelles or  
 CC microaggregates for inducing an immune response comprise: (a) a first  
 CC lipopeptide comprising at least one CTL (cytotoxic T-lymphocyte) epitope  
 CC and at least one lipid unit; and (b) a second lipopeptide comprising at  
 CC least one HTL (helper T-lymphocyte) epitope and at least one lipid unit  
 CC different from that of the first lipopeptide. This peptide represents  
 CC an example of a lipopeptide epitope used in the invention and is derived



PI Al-Garawi A, Kleanthous H, Miller C, Comen RP, Tomb J;  
 XX WPI: 1998-54293/46.  
 DR N-PSDB; AAX13994.  
 XX  
 PT New Isolated Helicobacter polynucleotides - used to develop products  
 PT for the diagnosis, prevention and treatment of Helicobacter  
 XX infections and gastrointestinal diseases  
 PS Claim 8; Page 301-306; 2054pp; English.  
 XX  
 CC This sequence represents a Helicobacter pylori GPO protein of the  
 CC invention. The polypeptides can be used for preventing or treating  
 CC Helicobacter infections, and gastroduodenal diseases associated with  
 CC these infections, including acute, chronic, and atrophic gastritis, and  
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be  
 CC used for the production of antibodies. The products can also be used for  
 CC detection and diagnosis.  
 CC  
 SQ Sequence 1230 AA;  
 XX  
 OY 5 TOEYKK 10  
 DB 173 tqevkk 178  
 XX  
 RESULT 11  
 ID AAY21365 standard; Protein; 7 AA.  
 AAY21365  
 AC AAY21365;  
 XX  
 DT 22-JUL-1999 (first entry)  
 XX  
 DE Human HUPF-1 mutant protein fragment 17.  
 XX  
 KW Human; beta-amyloid precursor protein; beta-APP; diagnosis: cancer;  
 KW frameshift mutation; age-related disease; neurodegenerative disorder;  
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;  
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;  
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;  
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;  
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;  
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;  
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;  
 KW high mobility group protein-C; neuroendocrine specific protein A.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 OS  
 XX  
 PN WO9845322-A2.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 02-APR-1998; 98WO-IB00705.  
 XX  
 PR 10-APR-1997; 97US-0043163.  
 XX  
 PA (UYUT-) RIJXSUNIV UTRECHT.  
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.  
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.  
 XX  
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;  
 XX WPI: 1998-609901/51.  
 DR N-PSDB; AAX75768.  
 XX  
 PT Diagnosing disease by detecting frameshift mutations in RNA or

PT corresponding protein mutations - used to diagnose cancer and  
 PT neurological diseases, particularly Alzheimer's disease, and also  
 PT for treatment and prevention with specific ribozymes or wild-type  
 PT RNA  
 XX  
 PS Disclosure; Figure 17; 258pp; English.  
 XX  
 CC This invention describes a novel method for the diagnosis of a disease  
 CC caused by, or associated with, an RNA molecule that has a frameshift  
 CC mutation. The method is used to diagnose age-related diseases, especially  
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's  
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,  
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II  
 CC and many others listed) or susceptibility to these disorders. The method  
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,  
 CC at an early stage. It is based on the observation that disease may be  
 CC caused by mutations in RNA rather than DNA. The invention describes the  
 CC use of neuronal system RNA molecules, specifically proteins including  
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated  
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule  
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,  
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic  
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma  
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group  
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.  
 CC  
 SQ Sequence 7 AA;  
 XX  
 OY 1 PVAPT 5  
 DB 2 pvapt 6  
 XX  
 RESULT 12  
 ID AAY53518 standard; Protein; 9 AA.  
 AAY53518  
 AC AAY53518;  
 XX  
 DT 18-JAN-2000 (first entry)  
 XX  
 DE HIV-1 P24 protein (aa 308-316) binds HLA-Cw.  
 XX  
 KW Lipopeptide; epitope; cytotoxic T lymphocyte; CTL; lipid; spacer; p53;  
 KW electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;  
 KW human immunodeficiency virus; hepatitis B virus; papilloma virus;  
 KW melanoma; malaria; parasite.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus type 1.  
 OS  
 XX  
 PN FR2776926-A1.  
 XX  
 PD 08-OCT-1999.  
 XX  
 PF 07-APR-1998; 98FR-0004323.  
 XX  
 PR 07-APR-1998; 98FR-0004323.  
 XX  
 PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.  
 PA (CNRS) CNRS CENT NAT RECH SCI.  
 PA (INSP) INST PASTEUR LILLE.  
 XX  
 PI Le Gal FA, Guillet JC, Gahery SH, Gras MH, Melnyk O, Tartar A;  
 XX WPI: 1999-583113/50.  
 DR  
 XX  
 PT New lipopeptide containing lipid regions and two epitopes, all

```

Db      1 pvpapq 6
      |||||
RESULT  8
ID      AAR45168 standard; Protein: 9 AA.
XX
AC      AAR45168;
XX
DT      16-JUN-1994 (first entry)
XX
DE      Listeria p60 peptide epitope.
XX
KW      Listeria monocytogenes; antibodies; immunoassay; conjugate.
XX
OS      Synthetic.
XX
PN      DE4318450-A.
XX
PD      16-DEC-1993.
XX
PE      03-JUN-1993; 93DE-4318450.
XX
PR      11-JUN-1992; 92DE-4219111.
XX
PR      25-NOV-1992; 92DE-4239567.
XX
PA      (MERE ) MERCK PATENT GMBH.
XX
PI      Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;
XX
PI      Hofmann G, Bubert A, Goebel W, Koehler S;
XX
DR      WPI: 1993-406956/51.
XX
PT      New primers for PCR detection of Listeria - including individual
PT      species, also new peptide(s) for raising antibodies for
PT      immunochemical detection
XX
PS      Disclosure: Fig 2; 19pp; German.
XX
CC      The sequence is that of a Listeria p60 peptide epitope which
CC      which may be used in the prodn. of antibodies for the detection
CC      of Listeria by immunoassay (partic. ELISA). It may be used as
CC      part of a method that allows determination of individual Listeria
CC      species, esp. L. monocytogenes.
XX
SQ      Sequence 9 AA:

Query Match      60.0%; Score 6; DB 14; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 pVAPRQ 6
      |||||
Db      4 pvpapq 9

RESULT  9
ID      AAR45159 standard; Protein: 20 AA.
XX
AC      AAR45159;
XX
DT      16-JUN-1994 (first entry)
XX
DE      Listeria p60 peptide epitope.
XX
KW      Listeria monocytogenes; antibodies; immunoassay; conjugate.
XX
OS      Synthetic.
XX
PA      Key
XX
PA      Location/Qualifiers

```

```

FT      Region      1..7
FT      /note= "opt. 0-7 amino acid residues"
FT      Region      14..20
FT      /note= "opt. 0-7 amino acid residues"
XX
XX      DE4318450-A.
XX
XX      16-DEC-1993.
XX
XX      03-JUN-1993; 93DE-4318450.
XX
XX      11-JUN-1992; 92DE-4219111.
XX
XX      25-NOV-1992; 92DE-4239567.
XX
XX      (MERE ) MERCK PATENT GMBH.
XX
XX      Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;
XX
XX      Hofmann G, Bubert A, Goebel W, Koehler S;
XX
XX      WPI: 1993-406956/51.
XX
XX      New primers for PCR detection of Listeria - including individual
XX      species, also new peptide(s) for raising antibodies for
XX      immunochemical detection
XX
XX      Claim 3; Page 4; 19pp; German.
XX
XX      The sequence is that of a Listeria p60 peptide epitope which
XX      which may be used in the prodn. of antibodies for the detection
XX      of Listeria by immunoassay (partic. ELISA). It may be used as
XX      part of a method that allows determination of individual Listeria
XX      species, esp. L. monocytogenes.
XX
XX      Sequence 20 AA:

Query Match      60.0%; Score 6; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 pVAPRQ 6
      |||||
Db      8 pvpapq 13

RESULT  10
ID      AAW98275 standard; Protein: 1230 AA.
XX
XX      AAW98275;
XX
XX      31-MAR-1999 (first entry)
XX
XX      H. pylori GHPo 690 protein.
XX
XX      GHPo protein; Helicobacter infection; gastroduodenal disease; gastritis;
XX      peptic ulcer disease.
XX
XX      Helicobacter pylori.
XX
XX      WO9843478-A1.
XX
XX      08-OCT-1998.
XX
XX      01-APR-1998; 98WO-US06371.
XX
XX      29-JUL-1997; 97US-0902615.
XX
XX      01-APR-1997; 97US-0833457.
XX
XX      24-JUN-1997; 97US-0881227.
XX
XX      (HUMA-) HUMAN GENOME SCI INC.
XX
XX      (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX

```

PA (SHAR/) SHARMA L R.  
 PA (VALS/) VAN ALSTYNE D.  
 XX  
 XX  
 PI Sharma LR, Van Alstyne D;  
 DR WPI: 1995-147431/19.  
 XX  
 XX New peptide(s) and corresp. antibodies for the treatment of  
 PT meningitis - the peptide(s) corresp. to homologous antigenic  
 PT sites on bacterial and viral agents and on chemokine(s), used for  
 PT detecting and preventing meningitis  
 PS Claim 47; Fig 7/10; 98pp; English.  
 XX  
 XX AAR73913 is the *Listeria monocytogenes* protein p60 precursor. It  
 CC contains the meningitis related antigenic sequences (MRHAS) claimed  
 CC in AAR73891-R73894 and AAR73903-R73906, which are recognised by a  
 CC monoclonal antibody from the hybridoma Rubella virus (RV)-1. The  
 CC claimed MRHAS peptides may be used in immunoassays to diagnose the  
 CC presence of bacterial and/or viral meningitis agents in a sample,  
 CC or in prophylactic and therapeutic meningitis treatments. The  
 CC peptides may also be used as vaccines against meningitis.  
 CC NB: Identified by matching corresponding MRHAS peptides.  
 XX  
 SQ Sequence 484 AA:

Query Match 100.0%; Score 10; DB 16; Length 484;  
 Best Local Similarity 100.0%; Pred. No. 0.0066;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPTOEVKK 10  
 |||||  
 DB 148 pvapltgevk 157

RESULT 6  
 AAR73903  
 ID AAR73903 standard; peptide; 7 AA.  
 XX  
 XX AAR73903:  
 XX  
 DT 05-DEC-1995 (first entry)  
 XX  
 DE *Listeria monocytogenes* protein p60 precursor peptide 151-157.  
 XX  
 KM *Listeria monocytogenes*; protein p60 precursor; vaccine;  
 KM meningitis related homologous antigenic sequence; MRHAS; RV-1;  
 KM immunoassay; diagnosis; treatment; prophylactic; bacterial;  
 KM viral; peptide 151-157.  
 XX  
 XX *Listeria monocytogenes*.  
 OS  
 XX  
 PN WO9509232-A.  
 PN  
 PD 06-APR-1995.  
 PD  
 XX  
 PF 28-SEP-1994; 94MO-CA00516.  
 PF  
 PR 28-SEP-1993; 93US-0127499.  
 PR  
 XX  
 PA (SHAR/) SHARMA L R.  
 PA (VALS/) VAN ALSTYNE D.  
 XX  
 XX Sharma LR, Van Alstyne D;  
 PI  
 XX WPI: 1995-147431/19.  
 DR  
 XX  
 PT New peptide(s) and corresp. antibodies for the treatment of  
 PT meningitis - the peptide(s) corresp. to homologous antigenic  
 PT sites on bacterial and viral agents and on chemokine(s), used for  
 PT detecting and preventing meningitis  
 XX

PS Claim 35; Page 76; 98pp; English.  
 XX  
 XX AAR73913 is the *Listeria monocytogenes* protein p60 precursor. It  
 CC contains the meningitis related antigenic sequences (MRHAS) claimed  
 CC in AAR73891-R73894 and AAR73903-R73906, which are recognised by a  
 CC monoclonal antibody from the hybridoma Rubella virus (RV)-1. The  
 CC claimed MRHAS peptides may be used in immunoassays to diagnose the  
 CC presence of bacterial and/or viral meningitis agents in a sample,  
 CC or in prophylactic and therapeutic meningitis treatments. The  
 CC peptides may also be used as vaccines against meningitis.  
 XX  
 SQ Sequence 7 AA:

Query Match 70.0%; Score 7; DB 16; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PTOEVKK 10  
 |||||  
 DB 1 ptgevk 7

RESULT 7  
 AAR54615  
 ID AAR54615 standard; Protein; 6 AA.  
 XX  
 XX AAR54615:  
 XX  
 DT 16-JUN-1994 (first entry)  
 XX  
 DE *Listeria p60* peptide epitope.  
 DE  
 XX  
 KM *Listeria monocytogenes*; antibodies; immunoassay; conjugate.  
 KM  
 XX  
 OS Synthetic.  
 OS  
 PN DE4318450-A.  
 PN  
 PD 16-DEC-1993.  
 PD  
 XX  
 PF 03-JUN-1993; 93DE-4318450.  
 PF  
 PR 11-JUN-1992; 92DE-4219111.  
 PR  
 PR 25-NOV-1992; 92DE-4239567.  
 XX  
 PA (MERE ) MERCK PATENT GMBH.  
 PA  
 XX Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;  
 PI Hofmann G, Bubert A, Goebel W, Koehler S;  
 PI  
 XX WPI: 1993-406956/51.  
 DR  
 XX  
 XX New primers for PCR detection of *Listeria* - including individual  
 PT species, also new peptide(s) for raising antibodies for  
 PT immunochemical detection  
 PT  
 XX  
 PS Disclosure: Page 3; 19pp; German.  
 PS  
 XX The sequence is that of a *Listeria p60* peptide epitope which  
 CC which may be used in the prodn. of antibodies for the detection  
 CC of *Listeria* by immunoassay (partic. ELISA). It may be used as  
 CC part of a method that allows determination of individual *Listeria*  
 CC species, esp. *L. monocytogenes*.  
 CC  
 XX  
 SQ Sequence 6 AA:

Query Match 60.0%; Score 6; DB 14; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPTO 6

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPRQEVKK 10  
 |||||  
 DB 3 PVAPRQEVKK 12

## RESULT 3

AAR73891  
 ID AAR73891 standard; peptide: 21 AA.

XX AAR73891;

XX 05-DEC-1995 (first entry)

DE Listeria monocytogenes protein p60 precursor peptide 144-164.

KW Listeria monocytogenes; protein p60 precursor; vaccine;

KW meningitis related homologous antigenic sequence; MRHAS; RV-1;

KW immunosassay; diagnosis; treatment; prophylactic; bacterial;

OS Listeria monocytogenes.

PN WO9509232-A.

XX 06-APR-1995.

PF 28-SEP-1994; 94WO-CA00516.

PR 28-SEP-1993; 93US-0127499.

PA (SHAR/) SHARMA L. R.

PI (VALS/) VAN ALSTYNE D.

PI Sharma LR, Van Alstyne D;

DR WPI: 1995-147431/19.

PT New peptide(s) and corresp. antibodies for the treatment of  
 PT meningitis - the peptide(s) corresp. to homologous antigenic  
 PT sites on bacterial and viral agents and on chemokine(s), used for  
 PT detecting and preventing meningitis

XX PS Claim 34; Page 74; 98pp; English.

XX AAR73913 is the Listeria monocytogenes protein p60 precursor. It  
 CC contains the meningitis related antigenic sequences (MRHAS) claimed  
 CC in AAR73891-R73894 and AAR73903-R73906, which are recognised by a  
 CC monoclonal antibody from the hybridoma Rubella virus (RV)-1. The  
 CC claimed MRHAS peptides may be used in immunoassays to diagnose the  
 CC presence of bacterial and/or viral meningitis agents in a sample,  
 CC or in prophylactic and therapeutic meningitis treatments. The  
 CC peptides may also be used as vaccines against meningitis.

XX SO Sequence 21 AA;

Query Match 100.0%; Score 10; DB 16; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 0.00045;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPRQEVKK 10  
 |||||  
 DB 5 PVAPRQEVKK 14

## RESULT 4

AAR45178  
 ID AAR45178 standard; protein: 478 AA.

XX AAR45178;

XX

DT 16-JUN-1994 (first entry)

XX Listeria p60 protein.

XX Immunogenic polypeptide; antibodies; immunoassay; conjugate.

XX Listeria monocytogenes.

PN DE4318450-A.

PD 16-DEC-1993.

PF 03-JUN-1993; 93DE-4318450.

PR 11-JUN-1992; 92DE-4239111.

PR 25-NOV-1992; 92DE-4239567.

XX (MERE ) MERCK PATENT GMBH.

PI Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;

PI Hofmann G, Buber A, Goebel W, Koehler S;

DR WPI: 1993-406956/51.

XX New primers for PCR detection of Listeria - including individual

PT species, also new peptide(s) for raising antibodies for

PT immunochemical detection

XX Disclosure; Fig 4; 19pp; German.

XX The sequence is that of the Listeria p60 protein. Antibodies

CC generated against the protein can be used in the detection

CC of Listeria by immunoassay (partic. ELISA). The detection method

CC allows determination of individual Listeria species, esp.

CC L. monocytogenes.

XX SO Sequence 478 AA;

Query Match 100.0%; Score 10; DB 14; Length 478;  
 Best Local Similarity 100.0%; Pred. No. 0.0065;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVAPRQEVKK 10  
 |||||  
 DB 148 PVAPRQEVKK 157

## RESULT 5

AAR73913  
 ID AAR73913 standard; protein: 484 AA.

XX AAR73913;

XX 05-DEC-1995 (first entry)

DE Listeria monocytogenes protein p60 precursor.

KW Listeria monocytogenes; protein p60 precursor; vaccine;

KW meningitis related homologous antigenic sequence; MRHAS; RV-1;

KW immunosassay; diagnosis; treatment; prophylactic; bacterial;

XX Listeria monocytogenes.

XX WO9509232-A.

XX 06-APR-1995.

XX 28-SEP-1994; 94WO-CA00516.

XX 28-SEP-1993; 93US-0127499.

961	4	40.0	124	16	AA075610	VH Fab M556-2 bind
962	4	40.0	124	16	AA075609	VH Fab H4H1-8 bind
963	4	40.0	124	16	AA075608	VH Fab H4H1-7 bind
964	4	40.0	124	16	AA075607	VH Fab H4H1-6 bind
965	4	40.0	124	16	AA075605	VH Fab H4H1-5 bind
966	4	40.0	124	16	AA075606	VH Fab H4H1-5 bind
967	4	40.0	124	16	AA075568	VH Fab M74 binds t
968	4	40.0	124	16	AA075569	VH Fab 3b1 binds t
969	4	40.0	124	16	AA075570	VH Fab 3b3 binds t
970	4	40.0	124	16	AA075572	VH Fab 3b9 binds t
971	4	40.0	124	16	AA075604	VH Fab H4H1-1 bind
972	4	40.0	124	16	AA075638	VH Fab H31 binds t
973	4	40.0	124	16	AA075639	VH Fab H33 binds t
974	4	40.0	124	16	AA075640	VH Fab H101 binds
975	4	40.0	124	16	AA075641	VH Fab H103 binds
976	4	40.0	124	17	AA001306	VH region of HIV n
977	4	40.0	124	17	AA001309	VH region of HIV n
978	4	40.0	124	17	AA001310	VH region of HIV n
979	4	40.0	124	17	AA001311	VH region of HIV n
980	4	40.0	124	17	AA001304	VH region of HIV n
981	4	40.0	124	17	AA001305	VH region of HIV n
982	4	40.0	124	17	AA001246	VH region of HIV n
983	4	40.0	124	17	AA001247	VH region of HIV n
984	4	40.0	124	17	AA001227	VH region of HIV n
985	4	40.0	124	18	AA033607	Human secreted pro
986	4	40.0	124	19	AA056732	Nucellus specific
987	4	40.0	124	20	AA074122	Human prostate tum
988	4	40.0	124	20	AA050146	Antibody F19 human
989	4	40.0	124	20	AA050147	Antibody F19 human
990	4	40.0	124	20	AA050148	Antibody F19 human
991	4	40.0	124	20	AA050149	Antibody F19 human
992	4	40.0	124	20	AA050150	Antibody F19 human
993	4	40.0	124	20	AA060074	Human endometrium
994	4	40.0	124	20	AA037522	Amino acid sequenc
995	4	40.0	124	21	AA010230	Human fetal kidney
996	4	40.0	124	21	AA037466	Arabidopsis thalia
997	4	40.0	124	21	AA095097	Anti-gp120 antibod
998	4	40.0	124	21	AA095098	Anti-gp120 antibod
999	4	40.0	124	21	AA095099	Anti-gp120 antibod
1000	4	40.0	124	21	AA095156	Anti-gp120 HC3 hea

## ALIGNMENTS

## RESULT 1

AA045173  
ID AA045173 standard; Protein: 10 AA.

XX	AA045173;	
AC	16-JUN-1994 (first entry)	
DT		
XX		
DE	Listeria p60 peptide epitope.	
XX		
KW	Listeria monocytogenes; antibodies; immunoassay; conjugate.	
XX		
OS	Synthetic.	
XX		
PN	DE0318450-A.	
XX		
PD	16-DEC-1993.	
XX		
PF	03-JUN-1993: 93DE-4318450.	
XX		
PR	11-JUN-1992: 92DE-4219111.	
XX		
PR	25-NOV-1992: 92DE-4239567.	
XX		
PA	(MERE ) MERCK PATENT GMBH.	
XX		
PI	Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;	
XX	Hofmann G, Bubert A, Goebel W, Koehler S;	
XX		

DR	WPI: 1993-406956/51.	
XX		
PT	New primers for PCR detection of Listeria - including individual	
PT	species; also new peptide(s) for raising antibodies for	
PT	immunochemical detection	
XX		
XX	Disclosure; Fig 2; 19pp; German.	
XX		
CC	The sequence is that of a Listeria p60 peptide epitope which	
CC	which may be used in the prodn. of antibodies for the detection	
CC	of Listeria by immunoassay (partic. ELISA). It may be used as	
CC	part of a method that allows determination of individual Listeria	
CC	species, esp. L. monocytogenes.	
XX		
SO	Sequence 10 AA;	

QY	1 PVAPTOEVKK 10	100.0%; Score 10; DB 14; Length 10;
		Best Local Similarity 100.0%; Pred. No. 0.00024;
Db	1 PVAPTOEVKK 10	Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 2

AA045172  
ID AA045172 standard; Protein: 12 AA.

XX	AA045172;	
AC	16-JUN-1994 (first entry)	
DT		
XX		
DE	Listeria p60 peptide epitope.	
XX		
KW	Listeria monocytogenes; antibodies; immunoassay; conjugate.	
XX		
OS	Synthetic.	
XX		
PN	DE0318450-A.	
XX		
PD	16-DEC-1993.	
XX		
PF	03-JUN-1993: 93DE-4318450.	
XX		
PR	11-JUN-1992: 92DE-4219111.	
XX		
PR	25-NOV-1992: 92DE-4239567.	
XX		
PA	(MERE ) MERCK PATENT GMBH.	
XX		
PI	Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;	
XX	Hofmann G, Bubert A, Goebel W, Koehler S;	
XX		
DR	WPI: 1993-406956/51.	
XX		
PT	New primers for PCR detection of Listeria - including individual	
PT	species; also new peptide(s) for raising antibodies for	
PT	immunochemical detection	
XX		
XX	Disclosure; Fig 2; 19pp; German.	
XX		
CC	The sequence is that of a Listeria p60 peptide epitope which	
CC	which may be used in the prodn. of antibodies for the detection	
CC	of Listeria by immunoassay (partic. ELISA). It may be used as	
CC	part of a method that allows determination of individual Listeria	
CC	species, esp. L. monocytogenes.	
XX		
SO	Sequence 12 AA;	

Query Match	100.0%; Score 10; DB 14; Length 12;
Best Local Similarity	100.0%; Pred. No. 0.00028;

815	4	40.0	117	22	AAB69685	Humanised M195 ant	888	4	40.0	121	17	AAR36611	Kaposi's sarcoma a
816	4	40.0	117	22	AAB69697	Humanised AF2 anti	889	4	40.0	121	17	AAR88504	VHmu for antibody
817	4	40.0	117	22	AAB69698	Human Eμ antibody	890	4	40.0	121	18	AAW13512	Anti-melanoma anti
818	4	40.0	118	13	AAR22569	Heavy chain VH15.4	891	4	40.0	121	19	AAW33584	Heavy chain of a h
819	4	40.0	118	13	AAR22421	Antibody produced	892	4	40.0	121	19	AAW33991	Anti-CD4 antibody
820	4	40.0	118	13	AAR28742	Heavy chain variab	893	4	40.0	121	20	AAV49236	N-terminal Cbpa se
821	4	40.0	118	14	AAR37611	hIL2R Ab H chain V	894	4	40.0	121	20	AAV32108	Choline binding pr
822	4	40.0	118	15	AAR61096	Human cancer cell	895	4	40.0	121	20	AAV37740	Chlamydia trachoma
823	4	40.0	118	15	AAR54758	Humanised HMFG1 he	896	4	40.0	121	21	AAW33900	Human colon cancer
824	4	40.0	118	15	AAR55126	Mouse-human chimr	897	4	40.0	121	21	AAW31682	Human secreted pro
825	4	40.0	118	15	AAR60305	Chimeric anti HIV	898	4	40.0	121	21	AAW14231	Arabisdopsis thalia
826	4	40.0	118	15	AAR63768	Humanised H chain	899	4	40.0	121	21	AAW18529	Zea mays protein f
827	4	40.0	118	16	AAR77184	Chimeric protein g	900	4	40.0	121	21	AAW33951	Arabisdopsis thalia
828	4	40.0	118	16	AAW27528	Heavy chain variab	901	4	40.0	121	21	AAW61696	Arabisdopsis thalia
829	4	40.0	118	18	AAW13516	Anti-melanoma anti	902	4	40.0	121	21	AAW00034	Human secreted pro
830	4	40.0	118	20	AAV30188	Amino acid sequenc	903	4	40.0	121	21	AAW00049	Human secreted pro
831	4	40.0	118	20	AAV47149	Humanised B9 anti	904	4	40.0	121	21	AAW00059	Human secreted pro
832	4	40.0	118	21	AAV71444	Humanised B9 anti	905	4	40.0	121	21	AAV32383	Pluripotent cell m
833	4	40.0	118	21	AAW02005	Human secreted pro	906	4	40.0	121	22	AAB69668	Humanised Fd38-80
834	4	40.0	118	21	AAV99558	Human Lh13 monoclo	907	4	40.0	122	10	AAW93697	Protein encoded by
835	4	40.0	118	21	AAV56679	Anti-integrin Mab	908	4	40.0	122	15	AAR60102	Mab heavy chain v1
836	4	40.0	118	21	AAV56681	Anti-erythropoiet	909	4	40.0	122	16	AAR75407	HIV-3 strain ANR70
837	4	40.0	119	13	AAR24560	Human x mouse modi	910	4	40.0	122	16	AAR76964	HSV-neutralising a
838	4	40.0	119	13	AAR24561	Human x mouse modi	911	4	40.0	122	19	AAW41087	Human matrix metel
839	4	40.0	119	13	AAR24562	Human x mouse modi	912	4	40.0	122	20	AAW90289	Human anti-GPIIb/I
840	4	40.0	119	13	AAR25730	Humanised VH regio	913	4	40.0	122	20	AAW90292	Human anti-GPIIb/I
841	4	40.0	119	14	AAR53364	N-terminally trunc	914	4	40.0	122	21	AAW03720	Human secreted pro
842	4	40.0	119	15	AAR55124	Mouse-human chimr	915	4	40.0	122	21	AAV70603	VH region of human
843	4	40.0	119	15	AAR60303	Chimeric anti HIV	916	4	40.0	122	22	AAB69691	Humanised CMV5 ant
844	4	40.0	119	16	AAR81324	Humanized VLA-4 an	917	4	40.0	123	16	AAR75571	VH Fab 3b4 binds t
845	4	40.0	119	16	AAR81325	Humanized VLA-4 an	918	4	40.0	123	17	AAR81323	Humanized VLA-4 an
846	4	40.0	119	16	AAR81331	Human 2°CL antibod	919	4	40.0	123	17	AAR92992	Humanised antibody
847	4	40.0	119	18	AAW22425	Humanised alpha-4	920	4	40.0	123	18	AAW22413	Humanised alpha-4
848	4	40.0	119	18	AAW22426	Anti-melanoma anti	921	4	40.0	123	18	AAW19881	CEA-specific anti
849	4	40.0	119	18	AAW13536	Anti-melanoma anti	922	4	40.0	123	18	AAW19886	CEA-specific anti
850	4	40.0	119	19	AAW49813	Amino acid sequenc	923	4	40.0	123	18	AAW19887	CEA-specific anti
851	4	40.0	119	19	AAW58503	Human heavy chain	924	4	40.0	123	18	AAW19888	CEA-specific anti
852	4	40.0	119	19	AAW53590	Humanised antibody	925	4	40.0	123	18	AAW19889	CEA-specific anti
853	4	40.0	119	19	AAW53591	Humanised antibody	926	4	40.0	123	19	AAW79228	Heavy chain variab
854	4	40.0	119	19	AAW53992	Anti-CD4 antibody	927	4	40.0	123	19	AAW76128	Human ICR-1.1 V-H
855	4	40.0	119	20	AAW95451	N-terminally trunc	928	4	40.0	123	19	AAW76125	Human ICR-8.1 V-H
856	4	40.0	119	21	AAW56018	Arabisdopsis thalia	929	4	40.0	123	19	AAW71259	Humanised murine a
857	4	40.0	119	21	AAV71241	Humanised antibody	930	4	40.0	123	19	AAW71255	Humanised murine a
858	4	40.0	119	21	AAW00045	Human secreted pro	931	4	40.0	123	20	AAV06264	Rest protein of CT
859	4	40.0	119	21	AAW03931	Human secreted pro	932	4	40.0	123	20	AAV00785	Antibody against I
860	4	40.0	119	21	AAW04052	Human secreted pro	933	4	40.0	123	20	AAV00782	Antibody against I
861	4	40.0	119	21	AAB08893	Human secreted pro	934	4	40.0	123	20	AAW81453	Humanised antibody
862	4	40.0	119	22	AAV56682	Humanized 369 heav	935	4	40.0	123	20	AAW81449	Humanised antibody
863	4	40.0	119	22	AAB20301	Arabisdopsis apopto	936	4	40.0	123	21	AAW13045	Protein sequence o
864	4	40.0	119	22	AAB69680	Humanised CMV5 ant	937	4	40.0	123	21	AAW83049	Humanised ICR-1.1
865	4	40.0	119	22	AAB69692	Human MO1 antibody	938	4	40.0	123	21	AAW82444	Humanised ICR-8.1
866	4	40.0	119	22	AAB62774	Human HIV-1 monocl	939	4	40.0	123	21	AAW82448	Humanised ICR-1.1
867	4	40.0	120	15	AAR47491	Humanised anti-CD1	940	4	40.0	123	21	AAV50920	Human fetal brain
868	4	40.0	120	15	AAR54795	SPA-reactive Igm h	941	4	40.0	123	21	AAV50752	Humanised murine a
869	4	40.0	120	15	AAR54796	SPA-reactive Igm h	942	4	40.0	123	21	AAV50756	Humanised murine a
870	4	40.0	120	17	AAW05825	Humanised M291 ant	943	4	40.0	123	21	AAV56714	Amino acid sequenc
871	4	40.0	120	18	AAW29998	Murine monoclonal	944	4	40.0	124	15	AAW45610	Monoclonal antibod
872	4	40.0	120	18	AAW27550	Human Ab heavy cha	945	4	40.0	124	15	AAR54330	Anti-HIV gp120 imm
873	4	40.0	120	18	AAW27551	Human Ab heavy cha	946	4	40.0	124	15	AAR54331	Anti-HIV gp120 imm
874	4	40.0	120	18	AAW27555	Human Ab heavy cha	947	4	40.0	124	15	AAR54332	Anti-HIV gp120 imm
875	4	40.0	120	18	AAW19882	CEA-specific antib	948	4	40.0	124	15	AAR54335	Anti-HIV gp120 imm
876	4	40.0	120	19	AAW44124	Heavy chain variab	949	4	40.0	124	15	AAR54336	Anti-HIV gp120 imm
877	4	40.0	120	20	AAW89350	Humanised 23F2c be	950	4	40.0	124	15	AAR54244	Anti-HIV gp120 imm
878	4	40.0	120	21	AAG27487	Arabisdopsis thalia	951	4	40.0	124	15	AAR54245	Anti-HIV gp120 imm
879	4	40.0	120	21	AAG00030	Human secreted pro	952	4	40.0	124	15	AAR54246	Anti-HIV gp120 imm
880	4	40.0	120	22	AAB47126	CDIFF-4, Incyte ID	953	4	40.0	124	15	AAR54269	Anti-HIV gp120 imm
881	4	40.0	120	22	AAB67381	P24B DNA, Human i	954	4	40.0	124	16	AAR75617	VH Fab M556-13 bin
882	4	40.0	120	22	AAB62747	Human HIV-1 monocl	955	4	40.0	124	16	AAR75616	VH Fab M556-5 bin
883	4	40.0	121	13	AAR25724	Humanised VH regio	956	4	40.0	124	16	AAR75615	VH Fab M556-16 bin
884	4	40.0	121	15	AAR54797	SPA-reactive Igm h	957	4	40.0	124	16	AAR75614	VH Fab M556-15 bin
885	4	40.0	121	16	AAR77874	Humanised mouse DR	958	4	40.0	124	16	AAR75613	VH Fab M556-7 bin
886	4	40.0	121	17	AAR99556	Fllea allergen Cten	959	4	40.0	124	16	AAR75612	VH Fab M556-3 bind
887	4	40.0	121	17	AAR97830	Kaposi's sarcoma a	960	4	40.0	124	16	AAR75611	VH Fab M556-3 bind

669	4	40.0	104	21	AAB42842	Human ORFX ORF2606	742	4	40.0	114	20	AAW95454	N-terminally trunc
670	4	40.0	104	21	AAV94933	Human secreted pro	743	4	40.0	114	21	AAV71550	Humanised antibody
671	4	40.0	105	19	AAV22856	SEQ ID NO. 52 from	744	4	40.0	114	21	AAV50974	Human FVII heavy
672	4	40.0	105	19	AAV22857	SEQ ID NO. 53 from	745	4	40.0	114	21	AAV75983	Murine skin cell p
673	4	40.0	105	19	AAV22858	SEQ ID NO. 54 from	746	4	40.0	114	21	AAV76033	Murine skin cell p
674	4	40.0	105	19	AAV22859	SEQ ID NO. 55 from	747	4	40.0	114	22	AAV55922	Skin cell protein,
675	4	40.0	105	19	AAW40579	Human lambda CL do	748	4	40.0	114	22	AAV55972	OP-1 based morphon
676	4	40.0	105	20	AAV38473	Human secreted pro	749	4	40.0	115	18	AAW12437	Rat quanylin prepr
677	4	40.0	105	20	AAV08746	Human lambda-CL do	750	4	40.0	115	20	AAV42461	Human phospholiest
678	4	40.0	105	20	AAV12366	Human 5' EST seque	751	4	40.0	115	20	AAV39287	Human Kabat subgro
679	4	40.0	105	20	AAW92426	Human lambda prote	752	4	40.0	115	20	AAW84095	Humanised HBV pre-
680	4	40.0	105	21	AAW27001	Human lambda CL do	753	4	40.0	115	21	AAW12169	Humanised HBV pre-
681	4	40.0	105	21	AAW00032	Human secreted pro	754	4	40.0	115	21	AAW12172	Human secreted pro
682	4	40.0	106	20	AAW01554	N-fermal CPBA se	755	4	40.0	115	21	AAW00036	Human secreted pro
683	4	40.0	106	20	AAV49228	Choline binding pr	756	4	40.0	115	21	AAW00051	Humanised VH regio
684	4	40.0	106	20	AAV32101	Amino acid sequenc	757	4	40.0	116	13	AAW25726	N-terminally trunc
685	4	40.0	106	20	AAW70803	Human pancreatic c	758	4	40.0	116	13	AAW53362	Humanised IL2 Mab
686	4	40.0	106	21	AAW54282	Human characterised hu	759	4	40.0	116	16	AAW66310	Humanised IL2 Mab
687	4	40.0	106	21	AAW23152	Human nucleic acid	760	4	40.0	116	17	AAW92218	Humanised IL2 Mab
688	4	40.0	106	21	AAW21030	Human secreted pro	761	4	40.0	116	17	AAW27698	Variable heavy cha
689	4	40.0	106	21	AAW03722	Human IgG1 lambda	762	4	40.0	116	18	AAW22418	Reshaped human AUK
690	4	40.0	106	21	AAV92193	Human secreted pro	763	4	40.0	116	19	AAW37812	Humanized anti-Tac
691	4	40.0	107	21	AAW51712	Human OREX ORF2304	764	4	40.0	116	19	AAW58514	Protein SEQ ID NO:
692	4	40.0	107	21	AAW42540	N. meningitidis BA	765	4	40.0	116	20	AAW95453	N-terminally trunc
693	4	40.0	107	21	AAW10493	Arbidopsis thaila	766	4	40.0	116	20	AAW18036	Eucalyptus grandis
694	4	40.0	107	21	AAW09358	Arbidopsis thaila	767	4	40.0	116	21	AAW18036	Eucalyptus grandis
695	4	40.0	107	21	AAW26325	Arbidopsis thaila	768	4	40.0	116	21	AAW18036	Eucalyptus grandis
696	4	40.0	107	21	AAW56019	Human secreted pro	769	4	40.0	116	21	AAW18036	Eucalyptus grandis
697	4	40.0	107	21	AAW03723	Human secreted pro	770	4	40.0	116	21	AAW18036	Eucalyptus grandis
698	4	40.0	107	21	AAV74672	Neisseria meningit	771	4	40.0	116	21	AAW00050	Zea mays protein f
699	4	40.0	107	21	AAV74673	Neisseria meningit	772	4	40.0	116	21	AAV50971	Human secreted pro
700	4	40.0	108	20	AAV05478	C. albicans Rb1 p	773	4	40.0	116	22	AAW69645	Human FVII antibo
701	4	40.0	108	21	AAW19160	Amino acid sequenc	774	4	40.0	116	22	AAW69646	Humanised anti-Tac
702	4	40.0	108	21	AAW08813	A human cullin-int	775	4	40.0	116	22	AAW69672	Humanised anti-Tac
703	4	40.0	108	21	AAW00033	Human secreted pro	776	4	40.0	116	22	AAW69672	Humanised anti-Tac
704	4	40.0	108	21	AAW03890	Human secreted pro	777	4	40.0	116	22	AAW69672	Humanised anti-Tac
705	4	40.0	109	18	AAW37348	Immunoglobulin G-1	778	4	40.0	117	13	AAW22358	Human antibody Eu
706	4	40.0	109	18	AAW41016	Anti-glutathione a	779	4	40.0	117	13	AAW25732	Protein encoded by
707	4	40.0	109	20	AAV49229	N-fermal CPBA se	780	4	40.0	117	14	AAW53363	Humanised VH regio
708	4	40.0	109	20	AAV32102	Choline binding pr	781	4	40.0	117	14	AAW36623	Human heavy chain
709	4	40.0	109	20	AAV28574	Secreted peptide c	782	4	40.0	117	15	AAW57476	CDR-grafted anti-R
710	4	40.0	109	21	AAW53666	Human colon cancer	783	4	40.0	117	15	AAW53329	Heavy chain variab
711	4	40.0	109	21	AAW58963	Arbidopsis thaila	784	4	40.0	117	15	AAW59514	Sequence of the ma
712	4	40.0	110	19	AAW55113	Streptococcus pneu	785	4	40.0	117	15	AAW59515	Sequence of the ma
713	4	40.0	110	20	AAW84099	Vitreonec tin alpha-	786	4	40.0	117	16	AAW66296	Human immunoglobul
714	4	40.0	110	21	AAW24425	Human PRO1274 prot	787	4	40.0	117	16	AAW66333	Human immunoglobul
715	4	40.0	110	21	AAW05562	Arbidopsis thaila	788	4	40.0	117	16	AAW66334	Human immunoglobul
716	4	40.0	110	21	AAW13654	Arbidopsis thaila	789	4	40.0	117	16	AAW66339	Human immunoglobul
717	4	40.0	110	21	AAW18889	Zea mays protein f	790	4	40.0	117	16	AAW66345	Human immunoglobul
718	4	40.0	110	21	AAV99380	Human PRO1274 (UNQ	791	4	40.0	117	16	AAW66316	Human immunoglobul
719	4	40.0	110	21	AAV87273	Human signal pepti	792	4	40.0	117	16	AAW66302	Human immunoglobul
720	4	40.0	110	22	AAW66129	Protein of the inv	793	4	40.0	117	16	AAW66311	Human immunoglobul
721	4	40.0	111	13	AAW22575	Heavy chain VH13.1	794	4	40.0	117	16	AAW66297	Human immunoglobul
722	4	40.0	111	21	AAW12314	Zea mays protein f	795	4	40.0	117	17	AAW03950	DNA fragment vH4.9.
723	4	40.0	111	21	AAW55547	Arbidopsis thaila	796	4	40.0	117	17	AAW92079	Mutine I108F VH CD
724	4	40.0	111	21	AAW03420	Human secreted pro	797	4	40.0	117	18	AAW41113	Human VHI family p
725	4	40.0	111	21	AAV92027	Human bone morphog	798	4	40.0	117	18	AAW24777	Human VHI gene VH4
726	4	40.0	111	22	AAW63325	Human breast cancer	799	4	40.0	117	19	AAW38656	S. pneumoniae dily
727	4	40.0	112	14	AAW43690	PBI.3/Humanised he	800	4	40.0	117	19	AAW58509	Protein SPO ID NO:
728	4	40.0	112	16	AAW66319	Human immunoglobul	801	4	40.0	117	20	AAW66146	Human bladder tumo
729	4	40.0	112	19	AAW77607	Staphylococcus aur	802	4	40.0	117	20	AAW40381	Amino acid sequenc
730	4	40.0	112	20	AAW49163	Soybean PITSURE pr	803	4	40.0	117	20	AAV37327	B. burgdorferi ant
731	4	40.0	112	21	AAW19019	Zea mays protein f	804	4	40.0	117	20	AAV19907	N-fermal CPBA se
732	4	40.0	112	22	AAW66407	Human Fab clone E9	805	4	40.0	117	20	AAW92252	BL172 polypeptide
733	4	40.0	112	22	AAW63407	Human breast cance	806	4	40.0	117	20	AAW95452	N-terminally trunc
734	4	40.0	113	17	AAW95650	Hepatitis C virus	807	4	40.0	117	20	AAW84097	Humanised anti-alp
735	4	40.0	113	17	AAW96561	Hepatitis C virus	808	4	40.0	117	20	AAW86141	Protein sequence o
736	4	40.0	113	17	AAW96562	Hepatitis C virus	809	4	40.0	117	21	AAW24952	Arbidopsis thaila
737	4	40.0	113	17	AAW96564	Hepatitis C virus	810	4	40.0	117	21	AAW09548	Human trypsin trunc
738	4	40.0	113	21	AAW40269	Human ORFX ORF33 p	811	4	40.0	117	21	AAW02814	Human trypsin trunc
739	4	40.0	113	22	AAW80629	Environmental stre	812	4	40.0	117	21	AAW92595	Trypsin truncated
740	4	40.0	114	14	AAW53361	N-terminally trunc	813	4	40.0	117	22	AAW69644	Human Eu antibody
741	4	40.0	114	20	AAW49136	Consensus sequence	814	4	40.0	117	22	AAW69650	Human Eu antibody

523	4	40.0	76	21	AAAG01671	Human secreted pro
524	4	40.0	76	22	AAAB5860	Murine TANGO 281 c
525	4	40.0	77	19	AAAB6001	S. pneumoniae derl
526	4	40.0	77	21	AAAB7219	Human secreted pro
527	4	40.0	78	21	AAAB52108	Human secreted pro
528	4	40.0	78	21	AAAB4691	Human 5' EST relat
529	4	40.0	78	22	AAAB87389	Human gene 48 enco
530	4	40.0	78	22	AAAB87412	Human gene 48 enco
531	4	40.0	79	19	AAAB22863	SEQ ID NO. 59 from
532	4	40.0	79	19	AAAB22864	SEQ ID NO. 60 from
533	4	40.0	79	20	AAAB73910	Human prostate tum
534	4	40.0	80	18	AAAB11223	S. pneumoniae indo
535	4	40.0	80	18	AAAB27747	Staphylococcus aur
536	4	40.0	80	21	AAAG00040	Human secreted pro
537	4	40.0	80	21	AAAG00067	Human secreted pro
538	4	40.0	81	20	AAAB28624	Human secreted pro
539	4	40.0	81	20	AAAB69936	Antigen from clust
540	4	40.0	81	21	AAAB73395	Human secreted pro
541	4	40.0	82	21	AAAG21847	Arabidopsis thalia
542	4	40.0	82	21	AAAB4976	Human 5' EST relat
543	4	40.0	84	21	AAAB03948	Human mesenchymal
544	4	40.0	85	12	AAAB14008	HIV-1 hxb2 gag 262
545	4	40.0	85	12	AAAB14458	HIV-1 hxb2 gag 262
546	4	40.0	85	21	AAAG02769	Human secreted pro
547	4	40.0	86	19	AAAB38421	Proteolipid protei
548	4	40.0	86	20	AAAB73904	Human prostate tum
549	4	40.0	86	20	AAAB73929	Human secreted pro
550	4	40.0	86	21	AAAG01952	Human secreted pro
551	4	40.0	86	22	AAAB35252	HIV-1 p24 protein
552	4	40.0	87	17	AAAB93003	Homologous sequenc
553	4	40.0	87	17	AAAB92991	Homologous sequenc
554	4	40.0	87	21	AAAG26156	Zea mays protein f
555	4	40.0	87	21	AAAB76169	Human secreted pro
556	4	40.0	88	13	AAAB22576	Heavy chain VH6.1
557	4	40.0	88	21	AAAG04407	Arabidopsis thalia
558	4	40.0	88	21	AAAG3200	Zea mays protein f
559	4	40.0	88	21	AAAG54249	Arabidopsis thalia
560	4	40.0	89	14	AAAB4271	Dp-74 VH gene hea
561	4	40.0	89	19	AAAB41088	Human matrix metal
562	4	40.0	89	19	AAAB41233	Alpha-v-beta-5 ant
563	4	40.0	89	21	AAAG04406	Arabidopsis thalia
564	4	40.0	89	21	AAAG54248	Arabidopsis thalia
565	4	40.0	90	21	AAAG26155	Zea mays protein f
566	4	40.0	91	21	AAAB42378	Human ORFX ORF2142
567	4	40.0	91	21	AAAG37140	Zea mays protein f
568	4	40.0	92	21	AAAB40314	Human ORFX ORF78 p
569	4	40.0	92	21	AAAG22778	Arabidopsis thalia
570	4	40.0	92	21	AAAG69934	Zea mays protein f
571	4	40.0	92	21	AAAG33234	Zea mays protein f
572	4	40.0	92	21	AAAG33943	Arabidopsis thalia
573	4	40.0	92	21	AAAG36731	Arabidopsis thalia
574	4	40.0	93	19	AAAB22860	SEQ ID NO. 56 from
575	4	40.0	93	19	AAAB75090	Human secreted pro
576	4	40.0	93	21	AAAG04405	Arabidopsis thalia
577	4	40.0	93	21	AAAG33473	Arabidopsis thalia
578	4	40.0	93	21	AAAG54247	Arabidopsis thalia
579	4	40.0	93	21	AAAG03293	Human secreted pro
580	4	40.0	94	19	AAAB22861	SEQ ID NO. 57 from
581	4	40.0	94	19	AAAB75151	Human secreted pro
582	4	40.0	94	19	AAAB75152	Human secreted pro
583	4	40.0	94	21	AAAB51698	Human secreted pro
584	4	40.0	94	21	AAAG00054	Human secreted pro
585	4	40.0	94	21	AAAG00065	Human secreted pro
586	4	40.0	94	22	AAAB54488	Human secreted pro
587	4	40.0	95	20	AAAB74161	Human prostate tum
588	4	40.0	95	20	AAAB60052	Human endometrium
589	4	40.0	95	21	AAAG14413	Arabidopsis thalia
590	4	40.0	95	21	AAAG57203	Arabidopsis thalia
591	4	40.0	95	21	AAAG59156	Arabidopsis thalia
592	4	40.0	95	21	AAAG00008	Human secreted pro
593	4	40.0	95	21	AAAB91621	Human secreted pro
594	4	40.0	96	20	AAAB67917	Human secreted pro
595	4	40.0	96	21	AAAG12504	Zea mays protein f
596	4	40.0	96	21	AAAG15817	Arabidopsis thalia
597	4	40.0	96	21	AAAG25016	Arabidopsis thalia
598	4	40.0	96	21	AAAG00039	Human secreted pro
599	4	40.0	96	21	AAAB56644	Partial peptide fr
600	4	40.0	96	21	AAAB56645	Partial peptide fr
601	4	40.0	96	21	AAAB56647	Partial peptide fr
602	4	40.0	96	21	AAAB56649	Partial peptide fr
603	4	40.0	97	14	AAAB53359	Osteogenic protein
604	4	40.0	97	14	AAAB53387	Murine osteogenic
605	4	40.0	97	15	AAAB57475	Human HV3 VH. Hom
606	4	40.0	97	17	AAAB92078	Human HV3 VH regio
607	4	40.0	97	20	AAAB27248	C. elegans CED-6 p
608	4	40.0	97	20	AAAB96966	Human osteogenic p
609	4	40.0	97	20	AAAB95444	Conserved 6 cystel
610	4	40.0	97	21	AAAG00068	Human secreted pro
611	4	40.0	97	21	AAAB55401	Human 5' EST relat
612	4	40.0	97	22	AAAB55852	Human TANGO 281 cy
613	4	40.0	98	10	AAAB95661	Human osteogenic p
614	4	40.0	98	14	AAAB33701	SP2 IgG VH segment
615	4	40.0	98	14	AAAB34272	Mutated human VH1
616	4	40.0	98	15	AAAB54800	SPA-reactive IgM h
617	4	40.0	98	16	AAAB72058	DP10 VH region. h
618	4	40.0	98	16	AAAB72059	HV1263 VH region.
619	4	40.0	98	16	AAAB72070	OE7H1.2 VH-1 H cha
620	4	40.0	98	19	AAAB54008	Human Anti-CD4 ant
621	4	40.0	98	21	AAAB40696	Human ORFX ORF460
622	4	40.0	98	21	AAAB50949	Human anti--factor
623	4	40.0	98	21	AAAB50951	Human anti--factor
624	4	40.0	98	21	AAAB50952	Human anti--factor
625	4	40.0	98	21	AAAB50954	Human anti--factor
626	4	40.0	98	21	AAAB50958	Human FVIII antibo
627	4	40.0	98	21	AAAB50970	Human FVIII antibo
628	4	40.0	99	20	AAAB95555	protein encoded by
629	4	40.0	99	21	AAAB18028	plins radicata myic
630	4	40.0	99	21	AAAB41554	Arabidopsis thalia
631	4	40.0	99	21	AAAG00046	Human secreted pro
632	4	40.0	99	21	AAAG00055	Human secreted pro
633	4	40.0	99	21	AAAG00066	Human secreted pro
634	4	40.0	99	21	AAAG00066	Rice glutamate 1-s
635	4	40.0	100	15	AAAB47238	Wild-type Feline H
636	4	40.0	100	17	AAAB03357	Human Immunodefice
637	4	40.0	100	21	AAAG00536	Human secreted pro
638	4	40.0	100	21	AAAG2052	Human secreted pro
639	4	40.0	100	21	AAAB1811	Streptococcus pneu
640	4	40.0	100	22	AAAB31970	Rice glutamate 1-s
641	4	40.0	101	20	AAAB35775	Chlamydia pneumoni
642	4	40.0	101	20	AAAB12050	Human 5' EST secre
643	4	40.0	101	21	AAAB54122	Human pancreatic c
644	4	40.0	102	13	AAAB95682	Human osteogenic p
645	4	40.0	102	13	AAAB25325	Hv region of human
646	4	40.0	102	14	AAAB53360	Osteogenic protein
647	4	40.0	102	18	AAAB36889	Human osteogenic p
648	4	40.0	102	18	AAAB36897	Mutant human osteo
649	4	40.0	102	18	AAAB36872	Human osteogenic p
650	4	40.0	102	18	AAAB56873	Mouse osteogenic p
651	4	40.0	102	18	AAAB18841	Human immunoglobuli
652	4	40.0	102	19	AAAB58889	S. pneumoniae derl
653	4	40.0	102	20	AAAB16708	WO9914235 Seq ID N
654	4	40.0	102	20	AAAB19744	SEQ ID NO 462 from
655	4	40.0	102	20	AAAB94855	Sequence of human.
656	4	40.0	102	20	AAAB89697	Human osteogenic p
657	4	40.0	102	20	AAAB95443	Conserved 7 cystel
658	4	40.0	102	21	AAAG00047	Human secreted pro
659	4	40.0	102	21	AAAG00071	Human secreted pro
660	4	40.0	102	21	AAAB09534	Human OP-1 C-termi
661	4	40.0	102	21	AAAB02800	Human OP-1 C-termi
662	4	40.0	102	21	AAAB92569	Human OP-1 amino a
663	4	40.0	103	18	AAAB41019	Anti-glutathione a
664	4	40.0	103	19	AAAB60484	Human epimorphin N
665	4	40.0	103	20	AAAB91335	Consensus sequence
666	4	40.0	103	20	AAAB91335	Consensus sequence
667	4	40.0	103	21	AAAB41148	Zea mays protein f
668	4	40.0	104	19	AAAG00061	Human secreted pro
	4	40.0	104	19	AAAB22878	SEQ ID NO. 74 from



377	4	40.0	35	21	AAV67809	Peptide #209 for d	450	4	40.0	56	21	AAV28299	Human secreted pro
378	4	40.0	37	18	AAW40369	Human breast cancer	451	4	40.0	56	21	AAAG01731	Human secreted pro
379	4	40.0	37	18	AAW30366	Fragment #2 of bon	452	4	40.0	56	21	AAV87131	Human secreted pro
380	4	40.0	37	20	AAV25880	Human secreted pro	453	4	40.0	56	22	AAV65854	Human TANGO 281 PS
381	4	40.0	37	21	AAV09516	Human OP-1 finger	454	4	40.0	56	22	AAV65862	Murine TANGO 281 P
382	4	40.0	37	21	AAV09559	Op-1 mutant protei	455	4	40.0	57	21	AAV54473	Zee mays protein f
383	4	40.0	37	21	AAV09562	Human OP-1 mutant	456	4	40.0	58	10	AAV01360	Amino acids 455-51
384	4	40.0	37	21	AAV09563	Human OP-1 mutant	457	4	40.0	58	10	AAV00062	Human secreted pro
385	4	40.0	37	21	AAV09565	Human OP-1 mutant	458	4	40.0	59	16	AAV73922	B. catarrhalis CD e
386	4	40.0	37	21	AAV09566	Human OP-1 mutant	459	4	40.0	59	21	AAV56233	Human secreted pro
387	4	40.0	37	21	AAV09573	Human OP-1 mutant	460	4	40.0	59	21	AAV59293	Arabidopsis thalia
388	4	40.0	37	21	AAV84032	Amino acid sequenc	461	4	40.0	59	21	AAV96790	Matize sucrose non-
389	4	40.0	38	14	AAV41075	HIV-1 gp120 C-term	462	4	40.0	60	18	AAV28031	Staphylococcus aur
390	4	40.0	38	20	AAV42830	Erythropoietin rec	463	4	40.0	60	18	AAV15568	PRN60. Feline leu
391	4	40.0	39	16	AAV66436	PCUS 4-18 (476-49	464	4	40.0	60	19	AAV20977	Human glial fibril
392	4	40.0	39	16	AAV66410	PCUS 4-18 (476-49	465	4	40.0	60	19	AAV72727	Proline rich pepti
393	4	40.0	39	18	AAV19025	HIV envelope glyco	466	4	40.0	60	22	AAV90561	Human secreted pro
394	4	40.0	39	20	AAV05347	HIV-1 CLUVAC pepi	467	4	40.0	61	20	AAV27507	E. coli beta'-subu
395	4	40.0	39	20	AAV05340	HIV-1 CLUVAC pepi	468	4	40.0	61	21	AAV02506	Human secreted pro
396	4	40.0	39	22	AAV59899	HIV-1 C5 region of	469	4	40.0	62	18	AAV27824	Staphylococcus aur
397	4	40.0	40	21	AAV05912	Human endothelial	470	4	40.0	62	18	AAV27825	Staphylococcus aur
398	4	40.0	40	21	AAV05913	Human endothelial	471	4	40.0	62	18	AAV03233	Human secreted pro
399	4	40.0	41	19	AAV22869	SEQ ID NO. 65 from	472	4	40.0	63	16	AAV76966	HSV-neutralising a
400	4	40.0	41	20	AAV42648	A. thaliana D22A73	473	4	40.0	63	16	AAV76967	HSV-neutralising a
401	4	40.0	41	20	AAV07893	Human secreted pro	474	4	40.0	63	16	AAV76965	HSV-neutralising a
402	4	40.0	42	9	AAV80686	Peptide 147 from t	475	4	40.0	63	16	AAV76971	HSV-neutralising a
403	4	40.0	42	18	AAV19055	SP10-BAL immunoge	476	4	40.0	63	16	AAV76972	HSV-neutralising a
404	4	40.0	42	19	AAV59179	Cela TMP-2 binding	477	4	40.0	63	16	AAV76973	HSV-neutralising a
405	4	40.0	42	20	AAV29924	Human CHD1 KRB do	478	4	40.0	63	21	AAV58469	Lung cancer associ
406	4	40.0	43	20	AAV91057	Flanking sequence	479	4	40.0	63	21	AAV51667	Human secreted pro
407	4	40.0	43	20	AAV91059	Flanking sequence	480	4	40.0	63	21	AAV18871	Zee mays protein f
408	4	40.0	43	20	AAV91063	Flanking sequence	481	4	40.0	63	21	AAV59038	Arabidopsis thalia
409	4	40.0	43	21	AAV27071	Beta-actenin flank	482	4	40.0	64	21	AAV55463	Arabidopsis thalia
410	4	40.0	43	21	AAV27073	Beta-actenin flank	483	4	40.0	65	19	AAV58662	A. pneumoniae derl
411	4	40.0	43	21	AAV27077	Beta-actenin flank	484	4	40.0	65	21	AAV38595	Human secreted pro
412	4	40.0	44	21	AAV76821	Cytoplasmic tail p	485	4	40.0	65	21	AAV76042	Human skin cell pr
413	4	40.0	44	21	AAV21848	Arabidopsis thalia	486	4	40.0	65	22	AAV55981	Arabidopsis thalia
414	4	40.0	45	10	AAV90818	Non-immunogenic am	487	4	40.0	66	21	AAV19536	Human colon cancer
415	4	40.0	45	20	AAV01143	Secreted protein e	488	4	40.0	66	21	AAV24624	Arabidopsis thalia
416	4	40.0	45	22	AAV87380	Human gene 39 enco	489	4	40.0	66	21	AAV34539	Arabidopsis thalia
417	4	40.0	46	19	AAV79183	Fusion immunoglob	490	4	40.0	66	21	AAV46116	Arabidopsis thalia
418	4	40.0	47	20	AAV36545	Fragment of human	491	4	40.0	66	21	AAV46138	Arabidopsis thalia
419	4	40.0	47	21	AAV83984	Human IL-18 bindin	492	4	40.0	67	20	AAV42644	Brassica napus D22
420	4	40.0	49	19	AAV46754	Exon trap L75917.	493	4	40.0	67	21	AAV53913	Human colon cancer
421	4	40.0	49	20	AAV25728	Human secreted pro	494	4	40.0	69	18	AAV32434	Mycobacterium tube
422	4	40.0	49	21	AAV56202	Human secreted pro	495	4	40.0	69	18	AAV32366	Mycobacterium tube
423	4	40.0	50	16	AAV71628	HIV gp120-C5. Syn	496	4	40.0	69	19	AAV81569	M. tuberculosis im
424	4	40.0	50	16	AAV29113	Enhanced CCK-B/gas	497	4	40.0	69	19	AAV64306	Mycobacterium tube
425	4	40.0	51	16	AAV77798	Transactivation do	498	4	40.0	69	20	AAV39108	M. tuberculosis an
426	4	40.0	51	16	AAV77799	Transactivation do	499	4	40.0	70	20	AAV38971	M. tuberculosis re
427	4	40.0	51	18	AAV29108	Enhanced CCK-B/gas	500	4	40.0	70	20	AAV93960	Human 53BP2:IP-2 p
428	4	40.0	51	21	AAV19538	Arabidopsis thalia	501	4	40.0	70	21	AAV53453	Human colon cancer
429	4	40.0	51	21	AAV24626	Arabidopsis thalia	502	4	40.0	70	21	AAV15847	Arabidopsis thalia
430	4	40.0	51	21	AAV34541	Arabidopsis thalia	503	4	40.0	71	21	AAV21088	Arabidopsis thalia
431	4	40.0	51	21	AAV46118	Arabidopsis thalia	504	4	40.0	72	16	AAV71350	Consensus sequence
432	4	40.0	51	21	AAV46140	Arabidopsis thalia	505	4	40.0	72	16	AAV71348	Mercury-binding re
433	4	40.0	51	21	AAV09225	Hepatitis GB virus	506	4	40.0	72	16	AAV71349	Mercury-binding re
434	4	40.0	52	10	AAV91361	Amino acids of the	507	4	40.0	72	19	AAV22887	SEQ ID NO. 83 from
435	4	40.0	52	21	AAV22870	White shrimp multi	508	4	40.0	72	21	AAV02885	Human secreted pro
436	4	40.0	52	21	AAV22971	White shrimp multi	509	4	40.0	73	22	AAV50534	Mouse C-CAM1 cyto
437	4	40.0	52	21	AAV66682	Membrane-bound pro	510	4	40.0	74	20	AAV03781	S. aureus polypept
438	4	40.0	52	22	AAV65205	Human PRO819 (UNC4	511	4	40.0	74	21	AAV56214	Human secreted pro
439	4	40.0	52	22	AAV50985	Human PRO819 prote	512	4	40.0	74	21	AAV37064	Arabidopsis thalia
440	4	40.0	52	22	AAV53093	Human angiotensins	513	4	40.0	74	21	AAV55678	Arabidopsis thalia
441	4	40.0	53	19	AAV22898	SEQ ID NO. 94 from	514	4	40.0	74	21	AAV60223	Arabidopsis thalia
442	4	40.0	54	19	AAV22867	SEQ ID NO. 63 from	515	4	40.0	74	22	AAV50533	Human C-CAM1 cyto
443	4	40.0	54	19	AAV22868	SEQ ID NO. 64 from	516	4	40.0	75	18	AAV28213	Staphylococcus aur
444	4	40.0	54	21	AAV59095	Breast and ovarian	517	4	40.0	75	20	AAV30854	Human secreted pro
445	4	40.0	54	21	AAV00997	Human secreted pro	518	4	40.0	75	20	AAV35810	Chlamydia pneumonia
446	4	40.0	55	8	AAV70360	Sequence of immuno	519	4	40.0	75	20	AAV12951	Amino acid sequenc
447	4	40.0	55	19	AAV22865	SEQ ID NO. 62 from	520	4	40.0	75	21	AAV18074	Arabidopsis thalia
448	4	40.0	55	19	AAV22865	SEQ ID NO. 61 from	521	4	40.0	76	20	AAV35859	Chlamydia pneumonia
449	4	40.0	55	20	AAV26920	Felv envelope glyc	522	4	40.0	76	20	AAV12175	Human 5' EST secre

231	4	40.0	11	14	AAR41809	M13TMDI mutated fr	304	4	40.0	22	14	AAR31965	BCR-132. Syntheti
232	4	40.0	11	17	AAR97523	Antigenic peptide,	305	4	40.0	22	14	AAR41293	peptide fragment F
233	4	40.0	11	19	AAW70126	Peptide produced b	306	4	40.0	22	19	AAW67032	peptide with N-ter
234	4	40.0	11	19	AAW50203	Peptide from a krl	307	4	40.0	22	20	AAV04034	HIV-1 p24 peptide
235	4	40.0	11	21	AAW22972	White shrimp p62 p	308	4	40.0	22	21	AAW12539	HIV-1 nucleoprotei
236	4	40.0	11	21	AAW09549	Peptide encoded by	309	4	40.0	22	13	AAW25941	N-terminal peptide
237	4	40.0	11	21	AAW02815	BMP mutant chimeri	310	4	40.0	23	14	AAW45164	Listeria p60 pepit
238	4	40.0	11	21	AAW92589	Peptide encoded by	311	4	40.0	23	14	AAW41294	Peptide fragment F
239	4	40.0	11	22	AAW53973	Human Kallikrein (	312	4	40.0	23	15	AAW41332	HIV gp120 epitope-
240	4	40.0	12	16	AAW83319	Kb-binding random	313	4	40.0	23	15	AAW63849	HIV-1 gp120-40 epi
241	4	40.0	12	16	AAW68778	Cytotoxic T lympho	314	4	40.0	23	16	AAW74046	Rice carbonic-anhy
242	4	40.0	12	21	AAW07094	Human peptido-mime	315	4	40.0	23	19	AAW50207	Peptide from a krl
243	4	40.0	13	17	AAW96827	Human snRNP-associ	316	4	40.0	23	19	AAW48108	HIV gp120 residues
244	4	40.0	13	19	AAW70128	Peptide produced b	317	4	40.0	23	20	AAW27500	E. coli beta'-subu
245	4	40.0	13	19	AAW50204	Peptide from a krl	318	4	40.0	24	8	AAW71151	Anti-HIV peptide 6
246	4	40.0	13	21	AAW22973	White shrimp p62 p	319	4	40.0	24	9	AAW80750	Sequence of peptid
247	4	40.0	14	19	AAW79191	Synthetic peptide	320	4	40.0	24	14	AAW33838	Cluster peptide PC
248	4	40.0	14	19	AAW56527	Antigenic HIV-1 pe	321	4	40.0	24	16	AAW6432	PCUS 4-18 (4/6-49
249	4	40.0	14	21	AAW22974	White shrimp p62 p	322	4	40.0	24	19	AAW54933	HIV gp120 envelope
250	4	40.0	14	21	AAW98509	Mouse WPI peptide	323	4	40.0	24	20	AAW05354	HIV-1 CUUVAC pepit
251	4	40.0	14	21	AAW98510	Human WPI peptide	324	4	40.0	24	21	AAW22976	White shrimp p62 p
252	4	40.0	15	10	AAW90224	Antigenic peptide-	325	4	40.0	25	7	AAW60708	HIV virus p18 prot
253	4	40.0	15	16	AAW79629	Endocarditis speci	326	4	40.0	25	11	AAW02316	Peptide derived fr
254	4	40.0	15	16	AAW79626	Endocarditis speci	327	4	40.0	25	13	AAW27336	Peptide corresp. t
255	4	40.0	15	18	AAW39011	Prostate Specific	328	4	40.0	25	16	AAW79982	Signal transductio
256	4	40.0	15	18	AAW30779	SH2 domain binding	329	4	40.0	25	19	AAW82523	HIV-1 p24 epitope
257	4	40.0	15	19	AAW78551	Human prostate spe	330	4	40.0	25	19	AAW82525	HIV-1 p24 epitope
258	4	40.0	15	19	AAW58037	Human prostate spe	331	4	40.0	25	20	AAW82525	HIV-1 p24 epitope
259	4	40.0	15	19	AAW58038	Human prostate spe	332	4	40.0	25	20	AAW81958	Human ENDO-II vari
260	4	40.0	15	19	AAW58039	Human prostate spe	333	4	40.0	26	13	AAW24867	Sequence of peptid
261	4	40.0	15	19	AAW58040	Human prostate spe	334	4	40.0	26	14	AAW30538	HIV discriminatory
262	4	40.0	15	20	AAW27174	Human SRT peptide	335	4	40.0	26	14	AAW30755	HIV discriminatory
263	4	40.0	15	21	AAW30407	Nuclear localisati	336	4	40.0	26	14	AAW41301	Peptide fragment F
264	4	40.0	15	21	AAW14221	Mutant HIV cleavag	337	4	40.0	26	14	AAW44111	HIV gp160 epitope.
265	4	40.0	15	21	AAW14223	PLC gamma SH3 doma	338	4	40.0	26	19	AAW76909	Fusion immunoglob
266	4	40.0	16	18	AAW25388	HTLV-1 envelope pr	339	4	40.0	26	19	AAW50208	Peptide from a krl
267	4	40.0	16	18	AAW09989	PRN4358. Feline I	340	4	40.0	26	21	AAW2977	White shrimp p62 p
268	4	40.0	16	18	AAW15570	Fusion immunoglob	341	4	40.0	27	10	AAW90819	Non-immunogenic am
269	4	40.0	16	19	AAW76907	Peptide from a krl	342	4	40.0	27	14	AAW30757	HIV discriminatory
270	4	40.0	16	19	AAW50206	Peptide from a krl	343	4	40.0	27	19	AAW79184	HIV gp120 C5 regio
271	4	40.0	16	19	AAW50205	Peptide from a krl	344	4	40.0	27	20	AAW79184	Human secreted pro
272	4	40.0	16	19	AAW37117	HTLV-I envelope an	345	4	40.0	28	15	AAW49685	Sequence of trypti
273	4	40.0	16	20	AAW00795	White shrimp p62 p	346	4	40.0	29	21	AAW33288	Pinus radiata tran
274	4	40.0	16	21	AAW22975	HTLV-I envelope pep	347	4	40.0	29	22	AAW49392	Simian immunodef
275	4	40.0	18	11	AAW02317	Peptide derived fr	348	4	40.0	29	22	AAW49393	Simian immunodef
276	4	40.0	18	19	AAW79190	Synthetic peptide	349	4	40.0	29	22	AAW49394	Simian immunodef
277	4	40.0	18	20	AAW94621	Src homology 3 dom	350	4	40.0	30	12	AAW15700	Human tumour-assoc
278	4	40.0	19	11	AAW01999	Peptide derived fr	351	4	40.0	30	13	AAW27045	Human heavy chain
279	4	40.0	19	11	AAW00940	Retroviral gag pep	352	4	40.0	30	16	AAW85163	Human ONS-M21 anti
280	4	40.0	19	14	AAW37695	gp120 epitope from	353	4	40.0	30	16	AAW70657	GLUT4 intracellular
281	4	40.0	19	16	AAW76724	Vasopressin type 2	354	4	40.0	30	16	AAW70658	Heavy chain framew
282	4	40.0	19	18	AAW33957	Cytotoxic T lympho	355	4	40.0	30	17	AAW66290	Fusion immunoglob
283	4	40.0	20	16	AAW68794	HIV peptide #37.	356	4	40.0	30	19	AAW76911	Framework I region
284	4	40.0	20	19	AAW08052	Fusion immunoglob	357	4	40.0	31	19	AAW79211	gp120 HTLV-III pep
285	4	40.0	20	19	AAW76943	Amino acid sequenc	358	4	40.0	31	18	AAW42189	Finger 2 domain of
286	4	40.0	20	21	AAW76875	Protein tyrosine p	359	4	40.0	31	19	AAW24744	HIV gp120 envelope
287	4	40.0	20	21	AAW93365	HIV-42-binding HIV	360	4	40.0	31	20	AAW48370	Human prostate can
288	4	40.0	20	21	AAW57262	Simian immunodef	361	4	40.0	31	20	AAW54936	Tomato Cu/Zn SOD e
289	4	40.0	20	22	AAW66445	Peptide ligand for	362	4	40.0	31	20	AAW82455	HIV-1 group O env
290	4	40.0	21	14	AAW49391	HTLV-I gp46 exten	363	4	40.0	31	21	AAW77377	HIV-1 group O env
291	4	40.0	21	14	AAW41614	Cladosporium herba	364	4	40.0	31	21	AAW77378	C-terminal portion
292	4	40.0	21	16	AAW34228	Conantokin-R (con-	365	4	40.0	32	18	AAW61490	Bacteriocalfitin-2
293	4	40.0	21	17	AAW87131	Conantokin-R (con-	366	4	40.0	32	18	AAW66676	WT1 derived immuno
294	4	40.0	21	17	AAW87131	Conantokin-R (con-	367	4	40.0	32	21	AAW98799	Hepatitis C virus
295	4	40.0	21	18	AAW94517	Peptide of the spe	368	4	40.0	33	16	AAW69686	Hepatitis C virus
296	4	40.0	21	18	AAW42769	Conantokin peptide	369	4	40.0	33	17	AAW89558	SE0 ID NO. 66 from
297	4	40.0	21	19	AAW49947	Conantokin peptide	370	4	40.0	34	19	AAW22870	Human p53 proline-
298	4	40.0	21	19	AAW48174	Peptide fragment o	371	4	40.0	34	19	AAW48443	Human prostate tum
299	4	40.0	21	20	AAW43600	Conantokin-R (con-	372	4	40.0	34	20	AAW74105	Human colon cancer
300	4	40.0	21	20	AAW86146	Synthetic peptide	373	4	40.0	34	21	AAW53925	Peptide produced b
301	4	40.0	22	11	AAW17925	Platelet derived g	374	4	40.0	35	19	AAW70114	Peptide #208 for d
302	4	40.0	22	14	AAW43929		375	4	40.0	35	21	AAW67806	
303	4	40.0	22	14	AAW30539		376	4	40.0	35	21	AAW67808	

85	5	50.0	399	18	AAW3417	Human integrin reg
86	5	50.0	399	20	AAV26927	Human integrin reg
87	5	50.0	404	19	AAW98621	H. pylori GHPO 245
88	5	50.0	421	20	AAV39368	SBWMS1 cytokine s
89	5	50.0	421	21	AAW12445	Human MSB1 protein
90	5	50.0	421	21	AAV53885	A suppressor of cy
91	5	50.0	422	22	AAW6464	Protein encoded by
92	5	50.0	428	20	AAV13450	Amino acid sequenc
93	5	50.0	445	19	AAW98793	H. pylori GHPO 125
94	5	50.0	445	19	AAW71558	Helicobacter polyp
95	5	50.0	452	18	AAW20598	H. pylori protein.
96	5	50.0	465	21	AAV81876	Streptococcus pneu
97	5	50.0	467	14	AAW41890	Bile acid sulphate
98	5	50.0	469	17	AAW97242	SCR phosphotyrosin
99	5	50.0	473	20	AAW98635	Protein encoded by
100	5	50.0	486	19	AAW37056	HIV-1 breakthrough
101	5	50.0	490	12	AAW15511	Sucrose phosphoryl
102	5	50.0	491	19	AAW37057	HIV-1 breakthrough
103	5	50.0	491	21	AAW69278	HIV-1 non-subtype
104	5	50.0	492	21	AAW69268	HIV-1 non-subtype
105	5	50.0	492	21	AAV96943	HIV synthetic gag
106	5	50.0	494	21	AAW69275	HIV-1 non-subtype
107	5	50.0	496	21	AAW69277	HIV-1 non-subtype
108	5	50.0	497	21	AAW69269	HIV-1 non-subtype
109	5	50.0	497	21	AAW69272	HIV-1 non-subtype
110	5	50.0	498	15	AAW62424	HIV-type virus MWP
111	5	50.0	498	15	AAW51691	HIV-type virus MWP
112	5	50.0	498	20	AAW93076	HIV isolate 5180 g
113	5	50.0	498	20	AAW93077	HIV isolate 5180 g
114	5	50.0	500	21	AAW69274	HIV-1 non-subtype
115	5	50.0	505	9	AAW80885	Sequence encoded b
116	5	50.0	505	19	AAW72992	HIV isolate LAV. MA
117	5	50.0	512	19	AAW68473	HIV-1 strain YBF30
118	5	50.0	515	10	AAW91235	(ENV-80)(GAG-VIT)
119	5	50.0	535	22	AAW49707	Small round struct
120	5	50.0	537	20	AAV34628	Protein Involved I
121	5	50.0	559	19	AAW50909	Alteromonas fortis
122	5	50.0	566	20	AAV49068	Amino acid sequenc
123	5	50.0	566	22	AAW31932	Amino acid sequenc
124	5	50.0	605	16	AAW9625	Endocarditis speci
125	5	50.0	612	19	AAW98516	H. pylori GHPO 130
126	5	50.0	649	20	AAW90005	Expressed antigen
127	5	50.0	655	14	AAW31041	srnk polypeptide.
128	5	50.0	662	20	AAV31940	Human rad17 cell c
129	5	50.0	663	20	AAV31941	Human rad17 cell c
130	5	50.0	669	20	AAV31939	Human rad17 cell c
131	5	50.0	670	17	AAW97869	Testis-associated
132	5	50.0	670	20	AAV31937	Human rad17 cell c
133	5	50.0	691	14	AAW38735	Beta-galactosidase
134	5	50.0	707	21	AAV95065	Candida albicans p
135	5	50.0	714	18	AAW35316	M. catarrhalis RI
136	5	50.0	730	9	AAW80618	Human Bone Morphog
137	5	50.0	730	9	AAW13669	C-proteinase encod
138	5	50.0	739	18	AAW25790	Gene 036 product d
139	5	50.0	762	20	AAV00939	Desulfurococcus M1
140	5	50.0	788	19	AAW75919	C-Proteinase sequen
141	5	50.0	797	21	AAV70963	Human Ras signalli
142	5	50.0	867	19	AAW98564	H. pylori GHPO 175
143	5	50.0	870	21	AAW10948	L. mexicana casein
144	5	50.0	986	18	AAW13670	C-proteinase encod
145	5	50.0	1006	21	AAW53126	Arbidopsis thalia
146	5	50.0	1013	19	AAW61539	Human cardiac/brai
147	5	50.0	1013	19	AAW40224	Human tollold-1like
148	5	50.0	1021	18	AAW23381	Allorrection assoc
149	5	50.0	1030	19	AAW53572	Human myosin light
150	5	50.0	1030	19	AAW41378	Human protein p164
151	5	50.0	1036	21	AAW31888	Arbidopsis thalia
152	5	50.0	1053	21	AAV99357	Human PRO1190 (UNQ
153	5	50.0	1053	22	AAW66106	Protein of the inv
154	5	50.0	1070	21	AAW31887	Arbidopsis thalia
155	5	50.0	1086	12	AAW11604	P450 17-alpha/P450
156	5	50.0	1095	21	AAW48592	Arbidopsis thalia
157	5	50.0	1130	21	AAW48591	Arbidopsis thalia
158	5	50.0	1187	16	AAW66451	Ar-4 protein (enco
159	5	50.0	1210	16	AAW66450	Ar-4 protein (enco
160	5	50.0	1245	13	AAW29029	Bacillus thuringie
161	5	50.0	1245	18	AAW16557	Bacillus thuringie
162	5	50.0	1245	18	AAW13872	B603a toxin. Bac
163	5	50.0	1245	19	AAW73105	B. t. toxin protein
164	5	50.0	1245	21	AAW13894	Bacillus thuringie
165	5	50.0	1245	22	AAW59884	8603 protein. Bac
166	5	50.0	1252	16	AAW80530	B. sphaerulic. Bac
167	5	50.0	1276	21	AAW31886	Arbidopsis thalia
168	5	50.0	1313	14	AAW36781	A3 maize Accase.
169	5	50.0	1336	21	AAW48590	Arbidopsis thalia
170	5	50.0	1409	21	AAW53125	Arbidopsis thalia
171	5	50.0	1420	21	AAW53124	Arbidopsis thalia
172	5	50.0	1484	20	AAW89721	Canine ribosome re
173	5	50.0	1566	16	AAW79643	Immunodominant ant
174	5	50.0	1627	12	AAW12789	M. pneumoniae p1 c
175	5	50.0	1627	15	AAW47911	Mycoplasma pneumon
176	5	50.0	1627	16	AAW67538	Cytadhesin P1. My
177	5	50.0	1743	19	AAW98879	H. pylori GHPO 175
178	5	50.0	2237	21	AAV58148	GURAY-3 polypeptel
179	5	50.0	2247	16	AAW27126	Equine rhinovirus
180	5	50.0	2325	17	AAW05590	Maize acetyl CoA c
181	5	50.0	2325	19	AAW56736	Maize ACCase enzym
182	5	50.0	2325	21	AAV44687	Maize acetyl CoA c
183	5	50.0	2476	20	AAW67738	P1g p105 zone pell
184	5	50.0	2625	21	AAW23832	Shewanella sp. SCR
185	5	50.0	2785	21	AAV57148	Human down-regulat
186	5	40.0	6	21	AAW09551	Peptide encoded by
187	5	40.0	6	21	AAW02817	BMP mutant chimeri
188	5	40.0	6	21	AAV92591	Peptide encoded by
189	5	40.0	6	22	AAW46417	Human IgG heavy ch
190	5	40.0	7	19	AAW48180	Conanotkin peptide
191	5	40.0	7	19	AAW49962	Conanotkin peptide
192	5	40.0	8	14	AAW43520	Sm B/B' epitope 16
193	5	40.0	8	17	AAW97521	Antigenic peptide,
194	5	40.0	8	20	AAV41615	Mammalian ion chan
195	5	40.0	8	21	AAV77393	HIV-1 group O env
196	5	40.0	9	14	AAW44262	Residues 68-76 of
197	5	40.0	9	14	AAW54620	Listeria p60 pepti
198	5	40.0	9	14	AAW42578	HIV envelope neutr
199	5	40.0	9	15	AAW57292	Bovine EK tryptic
200	5	40.0	9	15	AAW59259	Peptide fragment (
201	5	40.0	9	19	AAW79187	M38 murine monoclo
202	5	40.0	9	19	AAW70120	Peptide produced b
203	5	40.0	9	19	AAW78866	Prostate specific
204	5	40.0	9	20	AAV10315	T cell epitope/MHC
205	5	40.0	9	20	AAV10297	T cell epitope/MHC
206	5	40.0	9	20	AAW94619	Sic homology 3 dom
207	5	40.0	9	21	AAW45556	Human B99-1 HLA A3
208	5	40.0	9	21	AAW45557	Human B99-1 HLA A3
209	5	40.0	9	21	AAW33678	MHC class I associ
210	5	40.0	9	21	AAW98674	WT1 derived immuno
211	5	40.0	9	21	AAV98680	WT1 derived immuno
212	5	40.0	9	21	AAV98722	WT1 derived immuno
213	5	40.0	9	21	AAV98727	WT1 derived immuno
214	5	40.0	9	21	AAV98780	WT1 derived immuno
215	5	40.0	9	22	AAW76154	Tumour associated
216	5	40.0	10	17	AAW49590	Human leucocyte an
217	5	40.0	10	17	AAW49591	Human leucocyte an
218	5	40.0	10	17	AAW49592	Human leucocyte an
219	5	40.0	10	17	AAW97522	Antigenic peptide,
220	5	40.0	10	18	AAW40675	Peptide which bind
221	5	40.0	10	18	AAW42720	Antigenic decapt
222	5	40.0	10	18	AAW42721	Antigenic decapt
223	5	40.0	10	19	AAW76876	Fusion immunoglobu
224	5	40.0	10	19	AAW76878	Fusion immunoglobu
225	5	40.0	10	22	AAW75991	Human kallikrein (
226	5	40.0	10	22	AAW76201	Tumour associated
227	5	40.0	10	22	AAW76202	Tumour associated
228	5	40.0	11	10	AAW93053	HIV env protein an
229	5	40.0	11	14	AAW44263	Residues 68-78 of
230	5	40.0	11	14	AAW54627	Listeria innocua p

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:35:47 ; Search time 56.84 Seconds  
(without alignments)  
10.666 Million cell updates/sec

Title: US-09-372-036-31  
Perfect score: 10  
Sequence: 1 PVAPQTEVKK 10

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

1: /SIDSB/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SIDSB/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
22: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	ID	Description
1	100.0	10	AA45173	Listeria p60 pepti
2	100.0	12	AA45172	Listeria p60 pepti
3	100.0	21	AA43891	Listeria monocytog
4	100.0	478	AA45178	Listeria p60 prote
5	100.0	484	AA43913	Listeria monocytog
6	70.0	7	AA473903	Listeria monocytog
7	60.0	6	AA454615	Listeria p60 pepti
8	60.0	6	AA45168	Listeria p60 pepti
9	60.0	20	AA45159	Listeria p60 pepti
10	60.0	1230	AA498275	H. pylori GHPD 690
11	50.0	7	AA421365	Human HDPF-I mutan

12	50.0	9	AA453518	HIV-1 p24 protein
13	50.0	9	AA440367	Amino acid sequenc
14	50.0	20	AA426859	HIV-derived lipope
15	50.0	9	AA454624	Listeria innocua p
16	50.0	11	AA454624	Listeria innocua p
17	50.0	27	AA472848	Human p53 A76T mut
18	50.0	27	AA478130	Human secreted pro
19	50.0	87	AA427569	Human secreted pro
20	50.0	101	AA482898	CUB domain from BM
21	50.0	102	AA472875	Human p53 A76T/V12
22	50.0	107	AA475784	Human lymphocyte s
23	50.0	110	AA412211	Arabisidopsis thalia
24	50.0	110	AA459955	Arabisidopsis thalia
25	50.0	114	AA412210	Arabisidopsis thalia
26	50.0	121	AA459954	Arabisidopsis thalia
27	50.0	121	AA454455	Zea mays protein f
28	50.0	125	AA454454	Zea mays protein f
29	50.0	130	AA477657	Staphylococcus aur
30	50.0	132	AA453155	Chlamydia pneumoni
31	50.0	134	AA459953	Arabisidopsis thalia
32	50.0	135	AA436760	Amino acid sequenc
33	50.0	142	AA459953	Arabisidopsis thalia
34	50.0	142	AA453843	Arabisidopsis thalia
35	50.0	148	AA408988	Arabisidopsis thalia
36	50.0	148	AA453842	Arabisidopsis thalia
37	50.0	152	AA454453	Zea mays protein f
38	50.0	174	AA469904	Antigen from clust
39	50.0	179	AA455986	Arabisidopsis thalia
40	50.0	180	AA459942	Arabisidopsis thalia
41	50.0	180	AA455985	Arabisidopsis thalia
42	50.0	180	AA459941	Arabisidopsis thalia
43	50.0	181	AA425268	Eucalyptus grandis
44	50.0	201	AA434781	Chlamydia pneumoni
45	50.0	192	AA469639	Hepatitis C virus
46	50.0	192	AA469639	Hepatitis C virus
47	50.0	193	AA439866	HCT27 El protein.
48	50.0	196	AA455984	Arabisidopsis thalia
49	50.0	196	AA459940	Arabisidopsis thalia
50	50.0	201	AA469834	Protein encoded by
51	50.0	217	AA464638	Arabisidopsis thalia
52	50.0	221	AA460589	Human normal bladd
53	50.0	231	AA404036	HIV-1 (BHO clone)
54	50.0	240	AA496767	Z. mays MADS-box p
55	50.0	242	AA422390	Antigen ac-1b. El
56	50.0	247	AA493998	Human BR43x2, an l
57	50.0	251	AA408987	Arabisidopsis thalia
58	50.0	251	AA453841	Arabisidopsis thalia
59	50.0	256	AA4604637	Arabisidopsis thalia
60	50.0	274	AA421589	Arabisidopsis thalia
61	50.0	275	AA439342	Arabisidopsis thalia
62	50.0	280	AA4604636	Arabisidopsis thalia
63	50.0	288	AA479320	Corynebacterium gl
64	50.0	292	AA421588	Corynebacterium gl
65	50.0	293	AA475783	Arabisidopsis thalia
66	50.0	293	AA436312	Human lymphocyte s
67	50.0	293	AA494000	Human neutrophilic
68	50.0	293	AA494000	A transmembrane ac
69	50.0	293	AA439341	Arabisidopsis thalia
70	50.0	293	AA439341	Arabisidopsis thalia
71	50.0	303	AA428099	Human tumour necro
72	50.0	312	AA471914	Human lymphocyte s
73	50.0	312	AA471914	Human lymphocyte s
74	50.0	336	AA46021	Arabisidopsis thalia
75	50.0	343	AA416666	Human ORF1430
76	50.0	343	AA46020	Human ORF1430
77	50.0	347	AA460197	Maize DIMBOA biosy
78	50.0	351	AA460197	Arabisidopsis thalia
79	50.0	351	AA460197	Arabisidopsis thalia
80	50.0	367	AA460197	Arabisidopsis thalia
81	50.0	368	AA460197	Arabisidopsis thalia
82	50.0	371	AA429193	Amino acid sequenc
83	50.0	381	AA435327	Amino acid sequenc
84	50.0	388	AA404998	Mycobacterium spec
85	50.0	396	AA405432	Arabisidopsis thalia

**THIS PAGE BLANK (userio)**

---

Launis, Karen L.  
TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED  
NUCLEIC ACID CODING SEQUENCE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.  
STREET: 3054 Cornwalls Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,444A  
FILING DATE: 02-Jun-1995  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Melg9, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: S-18805/P1/CCCI577/CIP/DIV6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8587  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-08-459-444-19

Query Match 41.7%; Score 5; DB 3; Length 346;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 QEVK 12  
|||||  
DB 278 QEVK 282

RESULT 50  
US-08-137-627-4  
Sequence 4, Application US/08137627  
Patent No. 5681737  
GENERAL INFORMATION:  
APPLICANT: Gee, Pauline  
APPLICANT: Maron, Dorothy M.  
APPLICANT: Ames, Bruce N.  
TITLE OF INVENTION: A DETECTION SYSTEM FOR MUTAGENS THAT  
IDENTIFIES MUTAGENIC CHANGES  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Honbach, Test, Albritton & Herbert,  
ADDRESSEE: Attn: Karen S. Smith  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,627  
FILING DATE: 15-OCT-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Karen S.  
REGISTRATION NUMBER: 31,426  
REFERENCE/DOCKET NUMBER: A-58312/KSS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 359 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-137-627-4

Query Match 41.7%; Score 5; DB 1; Length 359;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STPVA 5  
|||||  
DB 249 STPVA 253

Search completed: August 15, 2001, 12:33:53  
Job time: 58 sec

APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttie, Janet L.  
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6018104artis Corporation  
STREET: Patent & Trademark Dept., 520 White Plains  
STREET: Rd., POB 2005  
CITY: Tarrytown  
STATE: New York  
COUNTRY: USA  
ZIP: 10591-9005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,595A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40403  
REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8582  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-459-595A-19

Query Match 41.7%; Score 5; DB 3; Length 346;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVKK 12  
|||||  
Db 278 QEVKK 282

RESULT 48  
US-08-459-504B-19  
Sequence 19, Application US/08459504B  
Patent No. 6075185  
GENERAL INFORMATION:  
APPLICANT: Kozziel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.

APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttie, Janet L.  
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6075185artis Corporation  
STREET: 3054 Cornwells Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,504B  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/459,595  
FILING DATE: 02-JUN-1995  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8587  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-459-504B-19

Query Match 41.7%; Score 5; DB 3; Length 346;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVKK 12  
|||||  
Db 278 QEVKK 282

RESULT 49  
US-08-459-444-19  
Sequence 19, Application US/08459444A  
Patent No. 6121014  
GENERAL INFORMATION:  
APPLICANT: Kozziel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.

APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttle, Janet L.  
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/951,715A  
FILING DATE: 25-SEP-1992  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Sprull, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8615  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-951-715A-19

Query Match 41.7%; Score 5; DB 1; Length 346;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 QEVK 12  
| | | | |  
DB 278 QEVK 282

RESULT 46  
US-08-459-448A-19  
Sequence 19, Application US/08459448A  
Patent No. 5859336  
GENERAL INFORMATION:  
APPLICANT: Kozziel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.

APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttle, Janet L.  
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 5859336artis Corporation  
STREET: Patent & Trademark Dept., 520 White Plains  
STREET: Rd., POB 2005  
CITY: Tarrytown  
STATE: New York  
COUNTRY: USA  
ZIP: 10591-9005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,448A  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40403  
REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8582  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 346 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-459-448A-19

Query Match 41.7%; Score 5; DB 2; Length 346;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 QEVK 12  
| | | | |  
DB 278 QEVK 282

RESULT 47  
US-08-459-595A-19  
Sequence 19, Application US/08459595A  
Patent No. 6018104  
GENERAL INFORMATION:  
APPLICANT: Kozziel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.



APPLICANT: PESCHER, PASCALE  
APPLICANT: ROMAÏN, FELIX  
TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM  
TITLE OF INVENTION: AND THEIR USES FOR VACCINES AND FOR THE DETECTION OF  
TITLE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 0660-0122-0 PCT  
CURRENT APPLICATION NUMBER: US/08/875,494  
CURRENT FILING DATE: 1997-08-01  
EARLIER APPLICATION NUMBER: PCT/FR96/00166  
EARLIER FILING DATE: 1996-02-01  
EARLIER APPLICATION NUMBER: 382184  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 325  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
US-08-875-494-2

Query Match 41.7%; Score 5; DB 4; Length 325;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
DB 66 TPVAP 70

RESULT 43  
US-08-444-818-152  
Sequence 152, Application US/08444818  
Patent No. 6150087  
GENERAL INFORMATION:  
APPLICANT: Chien, David Y.  
APPLICANT: Rutter, William J.  
TITLE OF INVENTION: NANOY Diagnostics and Vaccines  
NUMBER OF SEQUENCES: 777  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,818  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,590  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Harbin, Alisa A.  
REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885  
INFORMATION FOR SEQ ID NO: 152:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 339 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-444-818-152

Query Match 41.7%; Score 5; DB 4; Length 339;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
DB 113 PVAPT 117

RESULT 44  
US-08-444-818-156  
Sequence 156, Application US/08444818  
Patent No. 6150087  
GENERAL INFORMATION:  
APPLICANT: Chien, David Y.  
APPLICANT: Rutter, William J.  
TITLE OF INVENTION: NANOY Diagnostics and Vaccines  
NUMBER OF SEQUENCES: 777  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,818  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,590  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Harbin, Alisa A.  
REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885  
INFORMATION FOR SEQ ID NO: 156:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 339 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE: Human 27  
US-08-444-818-156

Query Match 41.7%; Score 5; DB 4; Length 339;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
DB 113 PVAPT 117

RESULT 45  
US-07-951-715A-19  
Sequence 19, Application US/07951715A  
Patent No. 5625136  
GENERAL INFORMATION:  
APPLICANT: Kosziel, Michael G.  
APPLICANT: Desai, Nalini M.

;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-382-184-2

Query Match 41.7%; Score 5; DB 1; Length 325;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPVAP 6  
|||||  
DB 66 TPVAP 70

RESULT 39  
US-08-641-356-2  
; Sequence 2, Application US/08641356  
; Patent No. 5866130  
; GENERAL INFORMATION:

; APPLICANT:

; TITLE OF INVENTION: MICROBACTERIAL PROTEINS,  
; TITLE OF INVENTION: MICROORGANISMS PRODUCING THEM AND THEIR USE FOR VACCINES  
; TITLE OF INVENTION: AND FOR THE DETECTION OF TUBERCULOSIS

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &

; STREET: 1755 S. Jefferson Davis Highway, Suite 400

; CITY: Arlington

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/641,356

; FILING DATE:

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/382,184

; FILING DATE: 01-FEB-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: OBLON, NORMAN F.

; REGISTRATION NUMBER: 24,618

; REFERENCE/DOCKET NUMBER: 22640720

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-413-3000

; TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 325 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-641-356-2

Query Match 41.7%; Score 5; DB 2; Length 325;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPVAP 6  
|||||  
DB 66 TPVAP 70

RESULT 40  
US-09-132-528-2  
; Sequence 2, Application US/09132528A  
; Patent No. 6221353

;; GENERAL INFORMATION:  
;; APPLICANT: LAOUEYRERIE, Anne  
;; APPLICANT: MARCHAL, Gilles  
;; APPLICANT: PESCHER, Pascale  
;; APPLICANT: ROMAIN, Felix

;; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM  
;; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF  
;; TITLE OF INVENTION: TUBERCULOSIS

;; FILE REFERENCE: 0660-0138-0DIY

;; CURRENT APPLICATION NUMBER: US/09/132,528A

;; CURRENT FILING DATE: 1998-08-11

;; EARLIER APPLICATION NUMBER: 08/641,356

;; EARLIER FILING DATE: 1996-04-30

;; NUMBER OF SEQ ID NOS: 5

;; SOFTWARE: Patentin Ver. 2.1

;; SEQ ID NO 2

;; LENGTH: 325

;; TYPE: PRT

;; ORGANISM: Mycobacterium tuberculosis

US-09-132-528-2

Query Match 41.7%; Score 5; DB 4; Length 325;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPVAP 6  
|||||  
DB 66 TPVAP 70

RESULT 41  
US-09-132-528-3  
; Sequence 3, Application US/09132528A  
; Patent No. 6221353

;; GENERAL INFORMATION:

;; APPLICANT: LAOUEYRERIE, Anne

;; APPLICANT: MARCHAL, Gilles

;; APPLICANT: PESCHER, Pascale

;; APPLICANT: ROMAIN, Felix

;; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM

;; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF

;; TITLE OF INVENTION: TUBERCULOSIS

;; FILE REFERENCE: 0660-0138-0DIY

;; CURRENT APPLICATION NUMBER: US/09/132,528A

;; CURRENT FILING DATE: 1998-08-11

;; EARLIER APPLICATION NUMBER: 08/641,356

;; EARLIER FILING DATE: 1996-04-30

;; NUMBER OF SEQ ID NOS: 5

;; SOFTWARE: Patentin Ver. 2.1

;; SEQ ID NO 3

;; LENGTH: 325

;; TYPE: PRT

;; ORGANISM: Mycobacterium tuberculosis

US-09-132-528-3

Query Match 41.7%; Score 5; DB 4; Length 325;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPVAP 6  
|||||  
DB 66 TPVAP 70

RESULT 42  
US-08-875-494-2  
; Sequence 2, Application US/08875494  
; Patent No. 6221610  
; GENERAL INFORMATION:  
; APPLICANT: LAOUEYRERIE, ANNE  
; APPLICANT: MARCHAL, GILLES

ORGANISM: Mycobacterium tuberculosis  
US-09-132-528-4

Query Match 41.7%; Score 5; DB 4; Length 286;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
DB 27 TPVAP 31

## RESULT 36

US-08-875-494-3  
Sequence 3, Application US/08875494  
Patent No. 6221610  
GENERAL INFORMATION:  
APPLICANT: LAQUEYERIE, ANNE  
APPLICANT: MARCHEL, GILLES  
APPLICANT: PESCHER, PASCALE  
APPLICANT: ROMAIN, FELIX  
TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM  
TITLE OF INVENTION: AND THEIR USES FOR VACCINES AND FOR THE DETECTION OF  
FILE REFERENCE: 0660-0122-0 PCT  
CURRENT APPLICATION NUMBER: US/08/875,494  
EARLIER FILING DATE: 1997-08-01 PCT/FR96/00166  
EARLIER APPLICATION NUMBER: 382184  
EARLIER FILING DATE: 1996-02-01  
EARLIER APPLICATION NUMBER: 382184  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 3  
LENGTH: 286  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
US-08-875-494-3

Query Match 41.7%; Score 5; DB 4; Length 286;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
DB 27 TPVAP 31

## RESULT 37

US-08-810-572A-2  
Sequence 2, Application US/08810572A  
Patent No. 5969102  
GENERAL INFORMATION:  
APPLICANT: Bram, Richard J.  
APPLICANT: von Bulow, Goltz  
TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS  
TITLE OF INVENTION: CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
STREET: Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/810,572A  
FILING DATE: 28-FEB-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1340-1-007  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 293 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-08-810-572A-2

Query Match 41.7%; Score 5; DB 2; Length 293;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APPOE 9  
DB 239 APPOE 243

## RESULT 38

US-08-382-184-2  
Sequence 2, Application US/08382184  
Patent No. 5714593  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: MICROBACTERIAL PROTEINS,  
TITLE OF INVENTION: MICROORGANISMS PRODUCING THEM AND THEIR USE FOR VACCINES  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER &  
ADDRESSEE: NEUSTADT, P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/382,184  
FILING DATE: 01-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 22640720  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 325 amino acids  
TYPE: amino acid

MOLECULE TYPE: protein  
US-08-468-544-12

Query Match 41.7%, Score 5, DB 4, Length 264;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TPVPA 5  
| | | | |  
Db 40 STVPA 44

RESULT 33  
US-08-382-184-3  
Sequence 3, Application US/08382184  
Patent No. 5714593  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: MICROBACTERIAL PROTEINS,  
TITLE OF INVENTION: MICROORGANISMS PRODUCING THEM AND THEIR USE FOR VACCINES  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
ADDRESS: NEUSTADT, P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/382,184  
FILING DATE: 01-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 22640720  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 286 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-08-382-184-3

Query Match 41.7%, Score 5, DB 1, Length 286;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 5: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 TPVAP 6  
| | | | |  
Db 27 TPVAP 31

RESULT 34  
US-08-641-356-3  
Sequence 3, Application US/08641356  
Patent No. 5866130  
GENERAL INFORMATION:  
APPLICANT:

TITLE OF INVENTION: MICROBACTERIAL PROTEINS,  
TITLE OF INVENTION: MICROORGANISMS PRODUCING THEM AND THEIR USE FOR VACCINES  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
ADDRESS: NEUSTADT, P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/641,356  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/382,184  
FILING DATE: 01-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 22640720  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 286 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
US-08-641-356-3

Query Match 41.7%, Score 5, DB 2, Length 286;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 5: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 TPVAP 6  
| | | | |  
Db 27 TPVAP 31

RESULT 35  
US-09-132-528-4  
Sequence 4, Application US/09132528A  
Patent No. 6221353  
GENERAL INFORMATION:  
APPLICANT: LAQUEYERRE, Anne  
APPLICANT: MARCHAL, Gilles  
APPLICANT: PESCHER, Pascale  
TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICORORGANISMS PRODUCING THEM  
TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF  
TITLE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 0660-0138-0DIY  
CURRENT APPLICATION NUMBER: US/09/132,528A  
CURRENT FILING DATE: 1998-08-11  
EARLIER APPLICATION NUMBER: 08/641,356  
EARLIER FILING DATE: 1996-04-30  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 4  
LENGTH: 286  
TYPE: PRT

ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,189  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/180,528  
FILING DATE:  
APPLICATION NUMBER: US/07/991,512  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MT 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ. ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-444-189-12

Query Match 41.7%; Score 5; DB 3; Length 264;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVA 5  
Db 40 STPVA 44

RESULT 31  
US-08-465-465-2  
Sequence 2, Application US/08465465  
Patent No. 6210968  
GENERAL INFORMATION:  
APPLICANT: Giulio Rattl  
TITLE OF INVENTION: RECOMBINANT CHLAMYDIA TRACHOMATIS PCP3  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 6210968rls  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch disk, 720 Kb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,465  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/229,980  
FILING DATE: 19-APR-1994

ATTORNEY/AGENT INFORMATION:  
NAME: Paul K. Legaard  
REGISTRATION NUMBER: 38,534  
REFERENCE/DOCKET NUMBER: CHIR-0039  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ. ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-465-465-2

Query Match 41.7%; Score 5; DB 4; Length 264;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVA 5  
Db 40 STPVA 44

RESULT 32  
US-08-468-544-12  
Sequence 12, Application US/08468544  
Patent No. 6248563  
GENERAL INFORMATION:  
APPLICANT: Rattl, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCID PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,544  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE: 07-FEB-1991  
APPLICATION NUMBER: IT MT 91A000314  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ. ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.  
TITLE OF INVENTION: GENETICALLY ENGINEERED COCCIDIOSIS  
SPOZOITE 21.5 KB ANTIGEN, AC-6B  
NUMBER OF SEQUENCES: 11  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/581,693  
FILING DATE: 12-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 215,162  
FILING DATE: 05-JUL-1988  
APPLICATION NUMBER: 746,520  
FILING DATE: 19-JUN-1985  
APPLICATION NUMBER: 627,811  
FILING DATE: 05-JUL-1984  
SEQ ID NO.: 3  
LENGTH: 242  
5273901-3

Query Match 41.7%; Score 5; DB 6; Length 242;  
Best Local Similarity 100.0%; Pred. No. 2,3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PTOEV 10  
|||||  
DB 36 PTOEV 40

RESULT 28  
5482709-2  
PATENT NO. 5482709  
APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON, SUSAN D.; POPE, SARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;  
AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.  
TITLE OF INVENTION: EIMERIA ANTIGENIC COMPOSITION WHICH  
ELICITS ANTIBODIES AGAINST AVIAN COCCIDIOSIS  
NUMBER OF SEQUENCES: 10  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/148,432  
FILING DATE: 08-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 581,693  
FILING DATE: 12-SEP-1990  
APPLICATION NUMBER: 215,162  
FILING DATE: 05-JUL-1989  
APPLICATION NUMBER: 746,520  
FILING DATE: 19-JUN-1985  
APPLICATION NUMBER: 627,811  
FILING DATE: 05-JUL-1984  
SEQ ID NO.: 2  
LENGTH: 242  
5482709-2

Query Match 41.7%; Score 5; DB 6; Length 242;  
Best Local Similarity 100.0%; Pred. No. 2,3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PTOEV 10  
|||||  
DB 36 PTOEV 40

RESULT 29  
US-08-969-644-12  
Sequence 12, Application US/08969644  
Patent No. 6096519  
GENERAL INFORMATION:  
APPLICANT: Rattl, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA

TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM: RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/969,644  
FILING DATE: 13-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/467,152  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO.: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-969-644-12

Query Match 41.7%; Score 5; DB 3; Length 264;  
Best Local Similarity 100.0%; Pred. No. 2,5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STPVA 5  
|||||  
DB 40 STPVA 44

RESULT 30  
US-08-444-189-12  
Sequence 12, Application US/08444189  
Patent No. 6110705  
GENERAL INFORMATION:  
APPLICANT: Rattl, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM: RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA

REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885  
INFORMATION FOR SEQ ID NO: 167:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: JH27  
US-08-444-818-167

Query Match 41.7%; Score 5; DB 4; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
|1111|  
DB 125 PVAPT 129

RESULT 25  
PCT-US95-10398-57  
Sequence 57, Application PC/TUS9510398  
GENERAL INFORMATION:  
APPLICANT: BORK, J., MILLER, R.H. AND  
APPLICANT: PURCELL, R.H.  
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND  
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS  
TITLE OF INVENTION: AND THE USE OF REGENITS DERIVED FROM THESE  
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES  
NUMBER OF SEQUENCES: 263  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/10398  
FILING DATE: 15-AUG-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/086,428  
FILING DATE: 29 JUNE 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/290/665  
FILING DATE: 15 AUGUST 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown

ORIGINAL SOURCE:  
ORGANISM: homoplans  
INDIVIDUAL ISOLATE: S18  
PCT-US95-10398-57

Query Match 41.7%; Score 5; DB 5; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
|1111|  
DB 50 PVAPT 54

RESULT 26  
US-08-969-721-8  
Sequence 8, Application US/08969721  
Patent No. 5972339  
GENERAL INFORMATION:  
APPLICANT: Walker, Bruce D.  
TITLE OF INVENTION: METHODS OF ELICITING ANTI-HIV-1  
TITLE OF INVENTION: HELPER T CELL RESPONSES  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: Windows 95  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/969,721  
FILING DATE: 13-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Ph.D., J.D., Janis K.  
REGISTRATION NUMBER: 34,819  
REFERENCE/DOCKET NUMBER: 08472/732001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-8906  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 231 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-969-721-8

Query Match 41.7%; Score 5; DB 2; Length 231;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEYK 11  
|1111|  
DB 178 TOEYK 182

RESULT 27  
5273901-3  
Patent No. 5273901  
APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,  
SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;

APPLICANT: BUKH, J., MILLER, R.H. AND  
APPLICANT: PURCELL, R.H.  
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE  
TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE  
TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN  
TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES  
NUMBER OF SEQUENCES: 159  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,570  
FILING DATE: 6-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/086,428  
FILING DATE: 29-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4070US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 751-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
ORIGINAL SOURCE:  
ORGANISM: homosapiens  
INDIVIDUAL ISOLATE: S18  
US-08-468-570-57

Query Match 41.7%; Score 5; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
Db 50 PVAPT 54

RESULT 23  
US-08-290-665A-57  
Sequence 57, Application US/08290665A  
Patent No. 5882852  
GENERAL INFORMATION:  
APPLICANT: BUKH, J., MILLER, R.H. AND  
APPLICANT: PURCELL, R.H.  
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND  
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS  
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE  
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES  
NUMBER OF SEQUENCES: 263  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE

CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290,665A  
FILING DATE: 15-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
ORIGINAL SOURCE:  
ORGANISM: homosapiens  
INDIVIDUAL ISOLATE: S18  
US-08-290-665A-57

Query Match 41.7%; Score 5; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
Db 50 PVAPT 54

RESULT 24  
US-08-444-818-167  
Sequence 167, Application US/08444818  
Patent No. 6150867  
GENERAL INFORMATION:  
APPLICANT: Chien, David Y.  
APPLICANT: Ruter, William J.  
TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
NUMBER OF SEQUENCES: 777  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chilton Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,818  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,590  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Harbin, Alisha A.  
REGISTRATION NUMBER: 33,895



PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,368  
FILING DATE:  
APPLICATION NUMBER: US 07/759,575  
FILING DATE: 13-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McClung, Barbara G.  
REGISTRATION NUMBER: 33,113  
REFERENCE/DOCKET NUMBER: 0205.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2708  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-542-44

Query Match 41.7%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
|||||  
DB 50 PVAPT 54

RESULT 20  
US-08-231-368-44  
Sequence 44, Application US/08231368  
Patent No. 5756312  
GENERAL INFORMATION:  
APPLICANT: Weiner, Amy J.  
APPLICANT: Houghton, Michael  
TITLE OF INVENTION: Immunoreactive polypeptide compositions  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chilton Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,368  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/759,575  
FILING DATE: 13-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McClung, Barbara G.  
REGISTRATION NUMBER: 33,113  
REFERENCE/DOCKET NUMBER: 0205.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2708  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-231-368-44

Query Match 41.7%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
|||||  
DB 50 PVAPT 54

RESULT 21  
US-08-440-210-44  
Sequence 44, Application US/08440210  
Patent No. 5766845  
GENERAL INFORMATION:  
APPLICANT: Weiner, Amy J.  
APPLICANT: Houghton, Michael  
TITLE OF INVENTION: Immunoreactive polypeptide compositions  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chilton Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,210  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,368  
FILING DATE:  
APPLICATION NUMBER: US 07/759,575  
FILING DATE: 13-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McClung, Barbara G.  
REGISTRATION NUMBER: 33,113  
REFERENCE/DOCKET NUMBER: 0205.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2708  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-210-44

Query Match 41.7%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
|||||  
DB 50 PVAPT 54

RESULT 22  
US-08-468-570-57  
Sequence 57, Application US/08468570  
Patent No. 5871962  
GENERAL INFORMATION:

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PROEV 10  
|||||

DB 43 PROEV 47

RESULT 17

US-08-086-428B-57

; Sequence 57, Application US/08086428B  
; Patent No. 5514339

; GENERAL INFORMATION:

; APPLICANT: BURKH, J., MILLER, R.H. AND

; APPLICANT: PURCELL, R.H.

; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED

; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE

; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE

; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN

; NUMBER OF SEQUENCES: 159

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/086.428B

; FILING DATE: 29-JUN-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: RICHARD W. BORK

; REGISTRATION NUMBER: 36,459

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 57:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 192 amino acids

; TYPE: amino acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; ORGANISM: homosapiens

; INDIVIDUAL ISOLATE: S18

US-08-086-428B-57

Query Match 41.7%; Score 5; DB 1; Length 192;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
|||||

DB 50 PVAPT 54

RESULT 18

US-08-440-103-44

; Sequence 44, Application US/08440103

; Patent No. 5670152

; GENERAL INFORMATION:

; APPLICANT: Weiner, Amy J.

; APPLICANT: Houghton, Michael

; TITLE OF INVENTION: Immunoreactive polypeptide compositions

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation

; STREET: 4560 Horton Street

; CITY: Emeryville

; STATE: CA

; COUNTRY: USA

; ZIP: 94608

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/440.103

; FILING DATE: 12-MAY-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/231.368

; FILING DATE:

; APPLICATION NUMBER: US 07/759,575

; FILING DATE: 13-SEP-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: McClung, Barbara G.

; REGISTRATION NUMBER: 33,113

; REFERENCE/DOCKET NUMBER: 0205.001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (510) 601-2708

; TELEFAX: (510) 655-3542

; INFORMATION FOR SEQ ID NO: 44:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 192 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-440-103-44

Query Match 41.7%; Score 5; DB 1; Length 192;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
|||||

DB 50 PVAPT 54

RESULT 19

US-08-440-542-44

; Sequence 44, Application US/08440542

; Patent No. 5670153

; GENERAL INFORMATION:

; APPLICANT: Weiner, Amy J.

; APPLICANT: Houghton, Michael

; TITLE OF INVENTION: Immunoreactive polypeptide compositions

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation

; STREET: 4560 Horton Street

; CITY: Emeryville

; STATE: CA

; COUNTRY: USA

; ZIP: 94608

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/440.542

; FILING DATE: 12-MAY-1995

; CLASSIFICATION: 424

RESULT 14  
US-09-001-984C-55  
; Sequence 55, Application US/09001984C  
; Patent No. 6245331  
; GENERAL INFORMATION:  
; APPLICANT: Lael, Suman  
; APPLICANT: Zolla-Pazner, Susan  
; APPLICANT: Belisle, John T  
; TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
; FILE REFERENCE: NYU-011  
; CURRENT APPLICATION NUMBER: US/09/001.984C  
; CURRENT FILING DATE: 1997-12-31  
; PRIOR APPLICATION NUMBER: 60/034.003  
; PRIOR FILING DATE: 1996-12-31  
; NUMBER OF SEQ ID NOS: 106  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 55  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis strain H37Rv  
US-09-001-984C-55

Query Match 41.7%; Score 5; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TPVAP 6  
11111  
Db 1 TPVAP 5

RESULT 15  
US-08-810-572A-4  
; Sequence 4, Application US/08810572A  
; Patent No. 5969102  
; GENERAL INFORMATION:  
; APPLICANT: Bram, Richard J.  
; APPLICANT: von Bulow, Gotz  
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS  
; TITLE OF INVENTION: CAML NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; STREET: Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/810.572A  
; FILING DATE: 28-FEB-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1340-1-007  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 107 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: C-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
US-08-810-572A-4

Query Match 41.7%; Score 5; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APPOE 9  
11111  
Db 53 APPOE 57

RESULT 16  
US-08-497-025-10  
; Sequence 10, Application US/08497025  
; Patent No. 5646251  
; GENERAL INFORMATION:  
; APPLICANT: Ruegg, Curtis L.  
; APPLICANT: Rivas, Alberto  
; APPLICANT: Laus, Reiner  
; APPLICANT: Engleman, Edgar G.  
; TITLE OF INVENTION: ALLOREACTION-ASSOCIATED ANTIGEN (AAR):  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourile and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/497.025  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/149,212  
; FILING DATE: 05-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 5490A-219  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 118 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..118  
; OTHER INFORMATION: /note= "A variable-type  
; OTHER INFORMATION: immunoglobulin-type domain of ARAg-h-1."  
US-08-497-025-10

Query Match 41.7%; Score 5; DB 1; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STERAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1-7  
OTHER INFORMATION: /note= "SOME OR ALL Xaa AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 14-20  
OTHER INFORMATION: /note= "SOME OR ALL Xaa AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
US-08-456-670B-17

Query Match 50.0%; Score 6; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 PVAPTQ 8  
| | | | |  
| | | | |  
DB 8 PVAPTQ 13

RESULT 13  
US-08-456-670B-35  
Sequence 35, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAWELZIK, MARTINA  
APPLICANT: LINKEWIELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STERAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria innocua  
US-08-456-670B-35

Query Match 41.7%; Score 5; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 QEVKK 12  
| | | | |  
| | | | |  
DB 7 QEVKK 11

ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-26

Query Match 66.7%; Score 8; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STPVAPRO 8  
Db 2 STPVAPRO 9

RESULT 10  
US-08-127-499A-28  
Sequence 28, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-28

Query Match 58.3%; Score 7; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 PTOEVKK 12  
Db 1 PTOEVKK 7

RESULT 11  
US-08-482-847-28  
Sequence 28, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
NUMBER OF SEQUENCES: 40  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-28

Query Match 58.3%; Score 7; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PTOEVKK 12  
Db 1 PTOEVKK 7

RESULT 12  
US-08-456-670B-17  
Sequence 17, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINXWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA

FILED DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5390  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-26

Query Match 100.0%; Score 12; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 4e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 STVPAPTOEVKK 12  
|||||  
Db 146 STVPAPTOEVKK 157

RESULT 8  
US-08-456-670B-31  
Sequence 31, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456.670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 421911.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:

NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
MOLECULE TYPE: linear  
TOPOLOGY: linear  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-31

Query Match 83.3%; Score 10; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PVAPPTQEVKK 12  
|||||  
Db 1 PVAPPTQEVKK 10

RESULT 9  
US-08-456-670B-26  
Sequence 26, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456.670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 421911.4  
FILING DATE: 11-JUN-1992

APPLICANT: LINKWELLER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 421911.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ. ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 478 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-40

Query Match 100.0%; Score 12; DB 2; Length 478;  
Best Local Similarity 100.0%; Pred. No. 3.9e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STPVAPTOEVK 12  
DB 146 STPVAPTOEVK 157

RESULT 6  
US-08-127-499A-26  
Sequence 26, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra

TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30.  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ. ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-26

Query Match 100.0%; Score 12; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 4e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STPVAPTOEVK 12  
DB 146 STPVAPTOEVK 157

RESULT 7  
US-08-482-847-26  
Sequence 26, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499

Best Local Similarity 100.0%; Pred. No. 2.5e-06;  
Matches 12: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVAPTOEVKK 12  
|||||  
Db 3 STPVAPTOEVKK 14

## RESULT 3

US-08-482-847-27

Sequence 27, Application US/08482847

Patent No. 5556757

GENERAL INFORMATION:

APPLICANT: VAN ALSTYNE, Diane

APPLICANT: SHARMA, Lawrence Rajendra

TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR

TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR

NUMBER OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF

NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley &amp; Lardner

STREET: 3000 K Street, N.W., Suite 500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/482,847

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/127,499

FILING DATE: 28-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 51916/104/INBI

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

US-08-482-847-27

Query Match 100.0%; Score 12; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.5e-06;  
Matches 12: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVAPTOEVKK 12  
|||||  
Db 3 STPVAPTOEVKK 14

## RESULT 4

US-08-456-670B-39

Sequence 39, Application US/08456670B

Patent No. 5932415

GENERAL INFORMATION:

APPLICANT: SCHUBERT, PETER

APPLICANT: NEUMANN, SIEGFRIED

APPLICANT: PAMELZIK, MARTINA

APPLICANT: LINKWEILER, WINFRIED

APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:

ADDRESSEE: MILLER, WHITE, ZELANO &amp; BRANIGAN, P.C.

STREET: 2200 CLARENDON BLVD., SUITE 1400

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: US

ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,670B

FILING DATE: 01-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/412,227

FILING DATE: 27-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/075,248

FILING DATE: 11-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4239567.4

FILING DATE: 25-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4219111.4

FILING DATE: 11-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: HAMLET-KING, DIANA

REGISTRATION NUMBER: 33,302

REFERENCE/DOCKET NUMBER: MERCK 1694D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-243-6333

TELEFAX: 703-243-6410

TELEX: 64191

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 232 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Listeria monocytogenes

STRAIN: EGD

US-08-456-670B-39

Query Match 100.0%; Score 12; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 2.1e-05;  
Matches 12: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVAPTOEVKK 12  
|||||  
Db 77 STPVAPTOEVKK 88

## RESULT 5

US-08-456-670B-40

Sequence 40, Application US/08456670B

Patent No. 5932415

GENERAL INFORMATION:

APPLICANT: SCHUBERT, PETER

APPLICANT: NEUMANN, SIEGFRIED

APPLICANT: PAMELZIK, MARTINA



977 4 33.3 236 3 US-08-493-071-15 Sequence 15, Appl  
978 4 33.3 236 4 US-09-049-672A-7 Sequence 7, Appl  
979 4 33.3 236 4 US-09-049-672A-13 Sequence 13, Appl  
980 4 33.3 236 4 US-08-961-083-96 Sequence 96, Appl  
981 4 33.3 237 1 US-08-096-946-10 Sequence 10, Appl  
982 4 33.3 237 1 US-08-096-946-11 Sequence 11, Appl  
983 4 33.3 237 2 US-08-672-564-3 Sequence 3, Appl  
984 4 33.3 237 2 US-08-844-024-2 Sequence 2, Appl  
985 4 33.3 237 3 US-08-718-547-2 Sequence 1, Appl  
986 4 33.3 237 3 US-08-768-859A-16 Sequence 16, Appl  
987 4 33.3 237 3 US-08-768-859A-21 Sequence 21, Appl  
988 4 33.3 237 3 US-08-767-820A-16 Sequence 16, Appl  
989 4 33.3 237 3 US-08-767-820A-21 Sequence 21, Appl  
990 4 33.3 237 3 US-08-767-820A-16 Sequence 16, Appl  
991 4 33.3 237 3 US-08-622-046B-1 Sequence 1, Appl  
992 4 33.3 237 3 US-08-622-046B-1 Sequence 1, Appl  
993 4 33.3 237 3 US-08-622-046B-1 Sequence 1, Appl  
994 4 33.3 237 3 US-08-622-046B-12 Sequence 12, Appl  
995 4 33.3 237 4 US-08-944-483-37 Sequence 37, Appl  
996 4 33.3 237 4 US-08-944-483-38 Sequence 38, Appl  
997 4 33.3 237 4 US-09-100-264-1 Sequence 1, Appl  
998 4 33.3 237 4 US-09-100-264-3 Sequence 3, Appl  
999 4 33.3 237 4 US-09-100-264-12 Sequence 12, Appl  
1000 4 33.3 237 5 PCT-US94-07329-10 Sequence 10, Appl

## ALIGNMENTS

RESULT 1  
US-08-456-670B-30  
; Sequence 30, Application US/08456670B  
; Patent No. 5932415  
; GENERAL INFORMATION:  
; APPLICANT: SCHUBERT, PETER  
; APPLICANT: NEUMANN, STEFRIED  
; APPLICANT: PAMELZIK, MARTINA  
; APPLICANT: LINXWEILER, WINFRIED  
; APPLICANT: BUNGER, CHRISTA  
; APPLICANT: HOFMANN, GOTTFRIED  
; APPLICANT: BUBERT, ANDREAS  
; APPLICANT: GOEBEL, WERNER  
; APPLICANT: KOHLER, STEFAN  
; TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
; TITLE OF INVENTION: LISTERIAS  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MILDEN, WHITE, ZELANO & BRANIGAN, P.C.  
; STREET: 2200 CLARENDON BLVD., SUITE 1400  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: US  
; ZIP: 22201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456,670B  
; FILING DATE: 01-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: US 08/412,227  
; FILING DATE: 27-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/075,248  
; FILING DATE: 11-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 4239567.4  
; FILING DATE: 25-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 4219111.4

; FILING DATE: 11-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HAMLET-KING, DIANA  
; REGISTRATION NUMBER: 33,302  
; REFERENCE/DOCKET NUMBER: MERCK 1694D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-243-6333  
; TELEFAX: 703-243-6410  
; TELEX: 64191  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; ORGANISM: Listeria monocytogenes  
; STRAIN: EGD

Query Match 100.0%; Score 12; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STVPAPTOEVKK 12  
DB 1 STVPAPTOEVKK 12

RESULT 2  
US-08-127-499A-27  
; Sequence 27, Application US/08127499A  
; Patent No. 5510264  
; GENERAL INFORMATION:  
; APPLICANT: VAN ALSTYNE, Diane  
; APPLICANT: SHARMA, Lawrence Rajendra  
; TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
; TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/127,499A  
; FILING DATE: 28-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 51916/102/INBI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
US-08-127-499A-27

Query Match 100.0%; Score 12; DB 1; Length 21;

















This Page Blank (usp10)

RT Isolates from Guinea-Bissau (West-Africa)."  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ246878; CAB95055.1; -  
 DR InterPro: IPR000777; -  
 DR Pfam: PF00516; GPI20; 1.  
 FT NON\_TER 1  
 FT NON\_TER 108  
 SQ SEQUENCE 108 AA; 12547 MM; FC22486D50C03680 CRC64;

Query Match 41.7%; Score 5; DB 14; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVK 11  
 |||||  
 DB 72 TOEVK 76

RESULT 50  
 09UG25  
 ID 09UG25 PRELIMINARY; PRT; 112 AA.  
 AC 09UG25;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE HYPOTHEICAL 12.8 KDA PROTEIN.  
 GN DKR2P564A122.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRIN;  
 RA Duesterhoeft A., Lauber J., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL110269; CAB53708.1; -  
 DR InterPro: IPR001496; -  
 DR InterPro: IPR001680; -  
 DR SMART: SM00320; WD40; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 112 AA; 12790 MM; 76795A2E6D112B5 CRC64;

Query Match 41.7%; Score 5; DB 4; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PTQEV 10  
 |||||  
 DB 91 PTQEV 95

Search completed: August 15, 2001, 12:42:49  
 Job time: 534 sec

RL AIDS Res. Hum. Retroviruses 16:1909-1914(2000).  
 DR EMBL: AF245564; AAG40677.1; -  
 FT NON\_TER 1  
 KW Coats protein; Envelope protein; Glycoprotein; Polyprotein;  
 KM Transmembrane.  
 FT NON\_TER 88  
 SQ SEQUENCE 88 AA: 10146 MW: 501A630530DC4FEB CRC64;

Query Match  
 Best local Similarity 41.7%; Score 5; DB 14; Length 88;  
 Best local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
 DB 73 OEYK 77

## RESULT 46

09MJT4 PRELIMINARY; PRT; 91 AA.  
 AC 09MJT4;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GAG PROTEIN (FRAGMENT).  
 GN GAG.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TW396;  
 RA Chang R.S., Lin R.-Y., Twu S.-C.,  
 RA Lin H.-C., Lin R.-Y., Salmien M.O., Liao S.-K., Wu A.M.,  
 RT "Hiv type 1 env gene diversity detected in Taiwan."  
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: U73091; AAD0206.1; -  
 DR InterPro: IPR000721; -  
 DR Pfam: PF00607; gag\_p24; 1.  
 DR NON\_TER 1  
 FT NON\_TER 91  
 SQ SEQUENCE 91 AA: 10492 MW: 21A08BD402AF874 CRC64;

Query Match  
 Best local Similarity 41.7%; Score 5; DB 14; Length 91;  
 Best local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVK 11  
 DB 60 TOEVK 64

## RESULT 47

068302 PRELIMINARY; PRT; 95 AA.  
 AC 068302;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E1 (GP32) (GP35)]  
 DE (FRAGMENT).  
 GN EL.  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepatitis C virus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HCV-BB43;  
 RA Songvilal S., Kanistanon D.;  
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1 SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
 DR EMBL: U73396; AAA64862.1; -

DR InterPro: IPR002519; -  
 DR Pfam: PF01539; HCV\_env; 1.  
 KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 KM Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 95  
 SQ SEQUENCE 95 AA: 10305 MW: 86270AA69397533A CRC64;

Query Match  
 Best local Similarity 41.7%; Score 5; DB 14; Length 95;  
 Best local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
 DB 23 PVAPT 27

## RESULT 48

068301 PRELIMINARY; PRT; 102 AA.  
 AC 068301;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E1 (GP32) (GP35)]  
 DE (FRAGMENT).  
 GN EL.  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepatitis C virus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HCV-BB8;  
 RA Songvilal S., Kanistanon D.;  
 RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1 SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
 DR EMBL: U73395; AAA64861.1; -  
 DR InterPro: IPR002519; -  
 DR Pfam: PF01539; HCV\_env; 1.  
 DR Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
 KM Transmembrane.  
 FT NON\_TER 1  
 FT NON\_TER 102  
 SQ SEQUENCE 102 AA: 11040 MW: 2C560825E0AD043E CRC64;

Query Match  
 Best local Similarity 41.7%; Score 5; DB 14; Length 102;  
 Best local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
 DB 24 PVAPT 28

## RESULT 49

091E25 PRELIMINARY; PRT; 108 AA.  
 AC 091E25;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN GP105 (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 2.  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11709;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Parreira R., Esteves A., Piedade J., Veneno T., Canas-Ferreira F.;  
 RT "Intra-individual genetic variability of immunodeficiency virus type 2

RA Nauci A. Jr.;  
 RT "HIV-2 genetic variation and DNA load in asymptomatic carriers and  
 RL AIDS cases in Guinea-Bissau.";  
 DR Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF023894; AAC63160.1;  
 DR InterPro; IPR000777;  
 DR Pfam; PF00516; GP120; 1.  
 KW Envelope protein.  
 FT NON\_TER 1 1  
 FT 77 77  
 SQ SEQUENCE 77 AA: 9063 MW: 9970B030D95EC9EC CRC64;  
 Query Match 41.7%; Score 5; DB 14; Length 77;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 8 OEKK 12 \*  
 Db 59 OEKK 63  
 RESULT 43  
 Q17728 PRELIMINARY; PRT; 82 AA.  
 AC Q17728;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE C06E4.2 PROTEIN.  
 GN C06E4.2  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 CC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,  
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Du Z., Gatlung S.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Waterston R.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U41277; AAA82477.1;  
 SQ SEQUENCE 82 AA: 9058 MW: E6CAC722EBB125B6 CRC64;  
 Query Match 41.7%; Score 5; DB 5; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 4 VAPTO 8  
 Db 40 VAPTO 44  
 RESULT 44

P91335  
 ID P91335 PRELIMINARY; PRT; 88 AA.  
 AC P91335;  
 DT 01-MAR-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE COSMID F55C7.  
 GN F55C7.2  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 CC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,  
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Du Z., Le T.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U80436; AAC71107.1;  
 SQ SEQUENCE 88 AA: 9771 MW: 59FB61FED7C0CE3 CRC64;  
 Query Match 41.7%; Score 5; DB 5; Length 88;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 STPVA 5  
 Db 31 STPVA 35  
 RESULT 45  
 Q9DR28 PRELIMINARY; PRT; 88 AA.  
 ID Q9DR28  
 AC Q9DR28;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DD88472;  
 RX MEDLINE=21002580; PubMed=11118076;  
 RA Abebe A., Pollakis G., Fontanet A.L., Fisseha B., Tegbaru B.,  
 RA Kiliphu A., Tesfaye G., Negassa H., Cornelissen M., Goudsmid J.,  
 RA Rinke de Wit T.F.;  
 RT "Identification of a genetic sub-cluster of HIV-1 subtype C (C')  
 RT widespread in Ethiopia.";

RC STRAIN-WISTAR;  
 RA MEDLINE-85207604; PubMed-2581948;  
 RX Northmann W., Heisig M., Kunz D., Hehrlich P.C.;  
 RT "Molecular cloning of cDNA sequences for rat alpha 2-macroglobulin and  
 RT measurement of its transcription during experimental inflammation";  
 RL J. Biol. Chem. 260:6200-6205(1985).  
 DR EMBL: M84369; AAA41594.1; -  
 DR InterPro: IPR001599; -  
 DR Pfam: PF00207; A2M; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 23 23  
 SO SEQUENCE 23 AA; 2526 MW; E2D721FB1B23876 CRC64;

Query Match 41.7%; Score 5; DB 11; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APTQE 9  
 |||||  
 DB 19 APTQE 23

RESULT 39  
 O9GL38 PRELIMINARY; PRT; 52 AA.  
 AC O9GL38;  
 DT 01-MAR-2001 (T-EMBLrel. 16, Created)  
 DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)  
 DE CALPASTATIN (FRAGMENT).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chung H.Y., Davis M.E., Hines H.C.;  
 RT "PCR-SSCP analysis of the bovine calpastatin gene domain 1 region";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY008267; AAG23869.1; -  
 FT NON\_TER 1 1  
 FT NON\_TER 52 52  
 SO SEQUENCE 52 AA; 5749 MW; CBD5A7449AFDDA89 CRC64;

Query Match 41.7%; Score 5; DB 6; Length 52;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVK 11  
 |||||  
 DB 4 TOEVK 8

RESULT 40  
 O47006 PRELIMINARY; PRT; 62 AA.  
 AC O47006;  
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)  
 DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)  
 DT 01-AUG-1998 (T-EMBLrel. 07, Last annotation update)  
 DE FLAGELLIN (FRAGMENT).  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-045:H23 (ECRC H23);  
 RA Kwang J., Wilson R., Yang S., He T.;  
 RL Clin. Diagn. Lab. Immunol. 0:0-0(0).

DR EMBL: U57313; AAB01993.1; -  
 FT NON\_TER 1 1  
 FT NON\_TER 62 62  
 SO SEQUENCE 62 AA; 6217 MW; 6BF545C9F2134414 CRC64;

Query Match 41.7%; Score 5; DB 2; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
 |||||  
 DB 15 TPVAP 19

RESULT 41  
 O9RRH1 PRELIMINARY; PRT; 75 AA.  
 AC O9RRH1;  
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)  
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)  
 DE HYPOTHETICAL 7.7 KDA PROTEIN.  
 GN DR2520.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-R1;  
 RX MEDLINE-20036896; PubMed-10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Mofatt K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamacheyan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 RT radiodurans R1";  
 RL Science 286:1571-1577(1999).  
 DR EMBL: AE002081; AAF12064.1; -  
 DR TIGR: DR2520; -  
 KW Hypothetical protein.  
 SO SEQUENCE 75 AA; 7704 MW; E8E47B7BBAE33B0E CRC64;

Query Match 41.7%; Score 5; DB 2; Length 75;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 VAPQ 8  
 |||||  
 DB 24 VAPQ 28

RESULT 42  
 O91692 PRELIMINARY; PRT; 77 AA.  
 AC O91692;  
 DT 01-NOV-1998 (T-EMBLrel. 08, Created)  
 DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)  
 DT 01-MAY-2000 (T-EMBLrel. 13, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 GN ENV.  
 OS Human immunodeficiency virus type 2.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11709;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-6059;  
 RA Leitner T., Norrgren H., Marquina S., Aaby P., Melbye M.,  
 RA Poulsen A.G., Larsen O., Dias F., Escanilla D., Albert J.,

Query Match 50.0%; Score 6; DB 5; Length 1052;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STVPAP 6  
 |||||  
 DB 383 STVPAP 388

## RESULT 35

025772 PRELIMINARY; PRT; 1230 AA.  
 AC 025772;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE HYPOTHELICAL 131.8 KDA PROTEIN.  
 GN HP1157.  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
 OC Helicobacter.  
 OC NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=26695 / ATCC 700392;  
 RX MEDLINE=97394467; PubMed=9252185;  
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
 RA McInerney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.R.,  
 RA Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D., Kelley J.M.,  
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,  
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
 RA Venter J.C.;  
 RA "The complete genome sequence of the gastric pathogen Helicobacter  
 pylori.";  
 RT Nature 388:539-547(1997).  
 RL EMBL: AE000621; AAD08205.1. -  
 DR EMBL: AE000621; AAD08205.1. -  
 DR TIGR: HP1157; -  
 DR InterPro: IPR002718; -  
 DR Pfam: PF01856; HP\_OMP; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1230 AA; 131847 MW; 4BF1506F1E072410 CRC64;

Query Match 50.0%; Score 6; DB 2; Length 1230;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12  
 |||||  
 DB 173 TOEVKK 178

## RESULT 36

09LBC3 PRELIMINARY; PRT; 1653 AA.  
 AC 09LBC3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE PAA.  
 GN PAA.  
 OS Streptococcus citreus.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 OC NCBI\_TaxID=1333;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=E49;  
 RA Tamura H., Kato H.;  
 RT "Cell surface antigen I/II - Streptococcus citreus.";

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB042239; BAA95000.1; -  
 DR InterPro: IPR001899; -  
 DR InterPro: IPR002965; -  
 DR Pfam: PF00746; Gram\_pos\_anchor; 1.  
 DR PRINTS: PRO1217; PRICHEXTENS.  
 DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; UNKNOWN.1.  
 SQ SEQUENCE 1653 AA; 180862 MW; 93F28E215BF598DE CRC64;

Query Match 50.0%; Score 6; DB 2; Length 1653;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPTO 8  
 |||||  
 DB 955 PVAPTO 960

## RESULT 37

09PVZ2 PRELIMINARY; PRT; 2037 AA.  
 AC 09PVZ2;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2001 (TREMBLrel. 16, Last annotation update)  
 DE NUCLEOPORIN CAN.  
 GN CAN.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99384295; PubMed=10454574;  
 RA Askjaer P., Bachl A., Wilm M., Bischoff R., Weeks D.L., Ogniewski V.,  
 RA Ono M., Niehrs C., Kjems J., Matraj I.W., Forrester M.;  
 RT "Rangrp-regulated interactions of CRM1 with nucleoporins and a  
 shuttling DEAD-box helicase.";  
 RL Mol. Cell. Biol. 19:6276-6285(1999).  
 DR EMBL: AJ243889; CAB53357.1; -  
 DR InterPro: IPR000515; -  
 DR InterPro: IPR001680; -  
 DR PROSITE: PS00402; BPD\_TRANSF\_INN\_MEMBER; UNKNOWN.1.  
 DR SMART: SM00320; WD40; 1.  
 KW Porin.  
 SQ SEQUENCE 2037 AA; 208931 MW; 2250EA49A0435635 CRC64;

Query Match 50.0%; Score 6; DB 13; Length 2037;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STVPAP 6  
 |||||  
 DB 1052 STVPAP 1057

## RESULT 38

063334 PRELIMINARY; PRT; 23 AA.  
 AC 063334;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE ALPHA-2-MACROGLOBULIN (FRAGMENT).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.

OX NCBI\_TaxID=11963;  
 RN [1]  
 RP SEQUENCE OF 980-989 FROM N.A.  
 RX MEDLINE=88004420; PubMed=2820721;  
 RA Flugel R.M., Rethwilm A., Maurer B., Darai G.;  
 RT "Nucleotide sequence analysis of the env gene and its flanking regions  
 of the human spumaretrovirus reveals two novel genes.";  
 RL EMBO J. 6:2077-2084(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Maurer B., Bannert H., Rethwilm A., Darai G., Flugel R.M.;  
 RL (in) Bolognesi D.P. (eds.):  
 Human retroviruses, cancer and AIDS, pp.75-88, Alan Liss Inc.,  
 New York (1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91101282; PubMed=1846194;  
 RA Muranyi W., Flugel R.M.;  
 RT "Analysis of splicing patterns of human spumaretrovirus by polymerase  
 chain reaction reveals complex RNA structures.";  
 RL J. Virol. 65:727-735(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9135779; PubMed=1651600;  
 RA Lochelt M., Zentgraf H., Flugel R.M.;  
 RT "Construction of an infectious DNA clone of the full-length human  
 spumaretrovirus genome and mutagenesis of the bel 1 gene.";  
 RL Virology 184:43-54(1991).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93348264; PubMed=8394017;  
 RA Lochelt M., Muranyi W., Flugel R.M.;  
 RT "Human foamy virus genome possesses an internal, Bel-1-dependent and  
 functional promoter.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:7317-7321(1993).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95000932; PubMed=7917520;  
 RA Weissenberger J., Flugel R.M.;  
 RT "Identification and characterization of the Bel 3 protein of human  
 foamy virus.";  
 RL AIDS Res. Hum. Retroviruses 10:595-600(1994).  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95380275; PubMed=7544460;  
 RA Kogel D., Aboud M., Flugel R.M.;  
 RT "Mutational analysis of the reverse transcriptase and ribonuclease H  
 domains of the human foamy virus.";  
 RL Nucleic Acids Res. 23:2621-2625(1995).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96135216; PubMed=8551561;  
 RA Lochelt M., Flugel R.M.;  
 RT "The human foamy virus pol gene is expressed as a Pro-Pol polypeptide  
 and not as a Gag-Pol fusion protein.";  
 RL J. Virol. 70:1033-1040(1996).  
 RN [9]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97126112; PubMed=8971036;  
 RA Bodem J., Lochelt M., Winkler I., Flower R.P., Delius H.,  
 Flugel R.M.;  
 RT "Characterization of the spliced pol transcript of feline foamy virus:  
 the splice acceptor site of the pol transcript is located in gag of  
 foamy viruses.";  
 RL J. Virol. 70:9024-9027(1996).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RX Flugel R.M.;  
 RL EMBL: U21247; AAB48113.1; -  
 KM Envelope protein.  
 SO SEQUENCE 989 AA; 113890 MM; E0E8338CE44E0A5 CRC64;

Query Match 50.0%; Score 6; DB 14; Length 989;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 TOEVRK 12  
 Db 185 TOEVRK 190  
 RESULT 34  
 ID Q9V550  
 AC Q9V550; PRELIMINARY; PRT; 1052 AA.  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE CG8598 PROTEIN.  
 GN CG8598.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Ephydroidea; Diptera; Ephydroidea; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Diptera; Ephydroidea; Diptera; Brachycera; Muscomorpha;  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SRRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Geiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,  
 RA April J.F., Adayani A., An H.-J., Andrews-Plamkoc C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokslein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Giordis A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merklov G., Malshtina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclebo J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003559; AAF50579.1; -  
 DR FlyBase: FBgn0035766; CG8598.  
 DR InterPro: IPR001969; -  
 DR PROSITE: PS00141; ASP\_PROTEASE; UNKNOWN.1.  
 SO SEQUENCE 1052 AA; 117413 MM; 504AE81BB3B12580 CRC64;

```

OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Caligaro S.T., Boude M., Cribbs D.L., Bourbon H.M.;
RT "tarrans", a novel member of the trithorax-group of homeotic gene
RT activators, encodes two protein isoforms related to the human cell-
RT cycle regulator p34(SEI-1).";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF227231; AAF43017.1;
DR FlyBase; FBgn0040071; tara.
DR InterPro; IPR000104;
DR PRINTS; PR00308; ANTIFREZE1.
SQ SEQUENCE 916 AA; 96242 MW; A67D331E012FA39E CRC64;

Query Match 50.0%; Score 6; DB 5; Length 916;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7
DB 718 TPVAPT 723

RESULT 30
O9NHCO PRELIMINARY; PRT; 916 AA.
AC O9NHCO;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE TARA 1A ISOFORM.
CN TARA OR C68889.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Caligaro S.T., Boude M., Cribbs D.L., Bourbon H.M.;
RT "tarrans", a novel member of the trithorax-group of homeotic gene
RT activators from Drosophila, encodes two protein isoforms related to
RT the human cell-cycle regulator p34(SEI-1).";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF227231; AAF43019.1;
DR FlyBase; FBgn0040071; tara.
DR InterPro; IPR000104;
DR PRINTS; PR00308; ANTIFREZE1.
SQ SEQUENCE 916 AA; 96242 MW; B76C220F103B98EE CRC64;

Query Match 50.0%; Score 6; DB 5; Length 916;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7
DB 718 TPVAPT 723

RESULT 31
O98830 PRELIMINARY; PRT; 988 AA.
ID O98830;
AC O98830;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-FEB-1997 (TREMBLrel. 02, Last annotation update)
DE PROVIRAL DNA, HSRV1 DELETION VARIANT.
CN ENV.
OS Human foamy virus.
OC Viruses; Retroid viruses; Retroviridae; Spumavirus.

```

```

OX NCBI_TaxID=11641;
RN [1]
RP SEQUENCE FROM N.A.
RA Schmidt M., Herchenröder O., Heeney J.L., Retzlaff A.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Retzlaff A.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y07723; CAA68994.1;
DR EMBL; Y07724; CAA69000.1;
DR EMBL; Y07725; CAA69004.1;
DR SEQUENCE 988 AA; 113762 MW; 0E9E53BF7D7C6B CRC64;

Query Match 50.0%; Score 6; DB 14; Length 988;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12
DB 185 TOEVKK 190

RESULT 32
O87041 PRELIMINARY; PRT; 988 AA.
ID O87041;
AC O87041;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE SFVCPZ, COMPLETE GENOME.
CN ENV.
OS Simian foamy virus.
OC Viruses; Retroid viruses; Retroviridae; Spumavirus.
OX NCBI_TaxID=11642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SFVCPZ;
RX MEDLINE-94240804; PubMed-8184531;
RA Herchenröder O., Renne R., Loncar D., Cobb E.K., Murthy K.K.,
RA Schneider J., Mergia A., Luciw P.A.;
RT "Isolation, cloning, and sequencing of simian foamy viruses from
RT chimpanzees (SFVCPZ): high homology to human foamy virus (HFV).";
RL Virology 201:187-199(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SFVCPZ;
RA Luciw P.A.;
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; U04327; AAA19979.1;
DR SEQUENCE 988 AA; 113352 MW; 9A88951475B9C62 CRC64;

Query Match 50.0%; Score 6; DB 14; Length 988;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12
DB 185 TOEVKK 190

RESULT 33
P90288 PRELIMINARY; PRT; 989 AA.
ID P90288;
AC P90288;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE ENVELOPE PROTEIN.
CN ENV.
OS Human spumaretrovirus (Foamy virus).
OC Viruses; Retroid viruses; Retroviridae; Spumavirus.

```



OC Ephyrdoidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SRRAIN-BERKELEY;  
 RA MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sultion G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wen R.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshnell A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector A.C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasmann D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003712; AAF55288.1; -  
 DR FlyBase: FBgn0040071; tara.  
 DR InterPro: IPR000104; -  
 DR PRINTS: PR00308; ANTIFREEZE1.  
 SQ SEQUENCE 904 AA; 94862 MW; 084F49645F5A6998 CRC64;

Query Match 50.0%; Score 6; DB 5; Length 904;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7  
 |||||  
 Db 706 TPVAPT 711

RESULT 27  
 O9NHCI PRELIMINARY; PRT; 912 AA.  
 AC O9NHCI;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE TARA 1B ISOFORM.  
 GN TARA OR CG6889.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyridae; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC "taranis", a novel trithorax-group of homeotic gene activators, encodes  
 RT two protein isoforms related to the human cell-cycle regulator  
 RT p34(Sel-1).";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF227212; AAF43018.1; -  
 DR FlyBase: FBgn0040071; tara.  
 DR InterPro: IPR000104; -  
 DR PRINTS: PR00308; ANTIFREEZE1.  
 SQ SEQUENCE 912 AA; 95728 MW; 0DBA6C2FE6B8F3E CRC64;

Query Match 50.0%; Score 6; DB 5; Length 912;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7  
 |||||  
 Db 714 TPVAPT 719

RESULT 28  
 O9NHBI PRELIMINARY; PRT; 912 AA.  
 AC O9NHBI;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE TARA 1B ISOFORM.  
 GN TARA OR CG6889.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephyridae; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Calargo S.T., Boube M., Cribbs D.L., Bourbon H.M.;  
 RT "taranis", a novel member of the trithorax-group of homeotic gene  
 RT activators from *Drosophila*, encodes two protein isoforms related to  
 RT the human cell-cycle regulator p34(Sel-1).";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF227213; AAF43020.1; -  
 DR FlyBase: FBgn0040071; tara.  
 DR InterPro: IPR000104; -  
 DR PRINTS: PR00308; ANTIFREEZE1.  
 SQ SEQUENCE 912 AA; 95698 MW; 1CAB7D3EFF7F544E CRC64;

Query Match 50.0%; Score 6; DB 5; Length 912;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7  
 |||||  
 Db 714 TPVAPT 719

RESULT 29  
 O9NHCI PRELIMINARY; PRT; 916 AA.  
 AC O9NHCI;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE TARA 1A ISOFORM.  
 GN TARA OR CG6889.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

RL Science 287:2185-2195(2000).  
 DR EMBL: AE003760; AAF56694.1; -  
 DR FLYbase: FBgn0039507; CG3361.  
 SQ SEQUENCE 775 AA; 85777 MW; AA3D5C6D6F871160 CRC64;

Query Match 50.0%; Score 6; DB 5; Length 775;  
 Best Local Similarity 100.0%; Pred. No. 99;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TPVAP 6  
 |||||  
 Db 358 TPVAP 363

RESULT 24  
 O9UAG1 PRELIMINARY; PRT; 792 AA.  
 AC O9UAG1;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE SALK-5.  
 OS Ephydria fluviatilis.  
 CC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;  
 CC Haplosclerida; Spongiillidae; Ephydria.  
 OC NCBT\_TaxID=31330;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99332084; PubMed=10405173;  
 RA Suga H., Ono K., Miyata T.;  
 RT "Multiple TGF-beta receptor related genes in sponge and ancient gene  
 RT duplication before the parazoan-eumetazoan split.";  
 RL FEBS Lett 453:346-350(1999).  
 CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: AB026828; BAA82605.1; -  
 DR InterPro: IPR000719; -  
 DR Pfam: PF00069; PKinase; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 KW ATP-binding; Serine/threonine-protein Kinase; Transferase.  
 SQ SEQUENCE 792 AA; 85470 MW; CED1CF0CDFCE523 CRC64;

Query Match 50.0%; Score 6; DB 5; Length 792;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 TPVAP 7  
 |||||  
 Db 575 TPVAP 580

RESULT 25  
 O9V48 PRELIMINARY; PRT; 819 AA.  
 AC O9V48;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CG16807 PROTEIN.  
 GN CG16807.  
 OS Drosophila melanogaster (Fruit Fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Eukaryota; Metazoa; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydriidae; Drosophilidae; Drosophila.  
 OC NCBT\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;

RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutcliffe G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.C.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Broxstein P., Brotlier P.,  
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davernport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spletter E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 CC -1- SIMILARITY: CONTAINS A RING (C3HC4-CLASS) ZINC FINGER.  
 DR EMBL: AE003527; AAF9474.1; -  
 DR FLYbase: FBgn0036621; CG16807.  
 DR InterPro: IPR000571; -  
 DR InterPro: IPR001841; -  
 DR Pfam: PF00097; zf-C3HC4; 1.  
 DR Pfam: PF00642; zf-CCCH; 1.  
 DR PROSITE: PS00518; ZINC\_FINGER\_C3HC4; 1.  
 DR SMART: SM00184; RING; 1.  
 KW Zinc-finger.  
 SQ SEQUENCE 819 AA; 90376 MW; 33D429076EDE7E2E CRC64;

Query Match 50.0%; Score 6; DB 5; Length 819;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 PVAPTO 8  
 |||||  
 Db 563 PVAPTO 568

RESULT 26  
 O9VEX7 PRELIMINARY; PRT; 904 AA.  
 AC O9VEX7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CG6889 PROTEIN.  
 GN TARA OR CG6889.  
 OS Drosophila melanogaster (Fruit Fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OY 1 STVPAP 6  
 DB 51 STVPAP 56

## RESULT 21

ID 022907 PRELIMINARY; PRT; 748 AA.

AC 022907;  
 DT 01-JAN-1998 (Tremblrel. 05, Created)  
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE RNA HELICASE ISOLG.  
 GN T08113.17.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;  
 RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,  
 RA Spriggs T.A., Mason T.M., Kerlavage A.R., Adams M.D., Somerville C.R.,  
 RA Venter J.C.;  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO DEAD/DEAD BOX HELICASE FAMILY.  
 CC -1- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES.  
 DR EMBL: AC002337; AAB63833.1; -;  
 DR EMBL: 26537; Atach; 257; 26537.  
 DR InterPro: IPR000629; -;  
 DR InterPro: IPR001410; -;  
 DR InterPro: IPR001650; -;  
 DR Pfam: PF00270; DEAD; 1.  
 DR Pfam: PF00271; helicase\_C.1.  
 DR PROSITE: PS00039; DEAD\_ATP\_HELICASE; 1.  
 DR SMART: SM00490; HELICG; 1.  
 KM ATP-binding; Helicase; RNA-binding.  
 SO SEQUENCE 748 AA; 81590 MW; 3F2FD0D5722D7364 CRC64;

Query Match 50.0%; Score 6; DB 10; Length 748;  
 Best Local Similarity 100.0%; Pred. No. 96;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPQ 8  
 DB 725 PVAPQ 730

## RESULT 22

ID 09HF24 PRELIMINARY; PRT; 750 AA.

AC 09HF24;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE REPERSED BY TUP1 PROTEIN 1.  
 GN RRT1.  
 OS Candida albicans (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5476;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA MEDLINE=20253093; Pubmed-10790384;  
 RA Braun B.R., Johnson A.D.;  
 RT "TUP1, CPH1 and EFG1 make independent contributions to filamentation  
 in candida albicans.";  
 RL Genetics 155:57-67(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=20435724; Pubmed-10978273;  
 RA Braun B.R., Head W.S., Wang M.X., Johnson A.D.;  
 RT "Identification and characterization of TUP1-regulated genes in  
 Candida albicans.";  
 RL Genetics 156:31-44(2000).  
 DR EMBL: AF254142; AAG09787.1; -;  
 SO SEQUENCE 750 AA; 76351 MW; 3B1C0C8EEAFCD72 CRC64;

Query Match 50.0%; Score 6; DB 3; Length 750;  
 Best Local Similarity 100.0%; Pred. No. 97;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STVPAP 6  
 DB 650 STVPAP 655

## RESULT 23

ID 09VB49 PRELIMINARY; PRT; 775 AA.

AC 09VB49;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)  
 DE CG3361 PROTEIN.

GN Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; Pubmed-10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokslein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,  
 RA Jaisli M., Kalush F., Kapran G.H., Ke Z., Kennison J.E., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshell A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacle J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";

DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
 DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)  
 DE POLYCYSTIC KIDNEY DISEASE-LIKE 2 PROTEIN.  
 GN PKD2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-20221369; PubMed-10756092;  
 RA Guo L., Schreiber T.H., Weremowicz S., Morton C.C., Lee C., Zhou J.;  
 RT "Identification and characterization of a novel polycystin family  
 member, polycystin-12, in mouse and human: sequence, expression,  
 alternative splicing, and chromosomal localization.";  
 RT Genomics 64:241-251(2000).  
 RL EMBL, AF182034; AAF65622.1; -;  
 DR InterPro: IPR000636; -;  
 DR InterPro: IPR001682; -;  
 DR InterPro: IPR002111; -;  
 SO SEQUENCE 624 AA; 73774 MW; 709FFA18542312E4 CRC64;

Query Match 50.0%; Score 6; DB 4; Length 624;  
 Best Local Similarity 100.0%; Pred. No. 83;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 TOEVKK 12  
 |||||  
 Db 298 TOEVKK 303

RESULT 19  
 008484 PRELIMINARY; PRT; 637 AA.  
 ID 008484  
 AC 008484; 000028;  
 DT 01-NOV-1996 (TRENBLREL. 01, Created)  
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
 DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)  
 DE CHROMOSOME XV READING FRAME ORF YOR070C.  
 OS YOR29-21 OR GYPI.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bohn C., Bolotin-Fukuhara M., Daignan-Fornier B., Dang D.V.,  
 RA Valens M.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MIRS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97279235; PubMed-9133743;  
 RA Valens M., Bohn C., Daignan-Fornier B., Dang V., Bolotin-Fukuhara M.;  
 RT "The sequence of a 54.7 kb fragment of yeast chromosome XV reveals the  
 presence of two tRNAs and 24 new open reading frames.";  
 RL Yeast 13:379-390(1997).  
 DR EMBL, Z70478; CA99263.1; -;  
 DR EMBL, 270678; CA94555.1; -;  
 DR SGD, S0005596; GYPI.  
 DR InterPro: IPR000195; -;  
 DR Pfam: PF00566; TBC; 1.  
 DR SMART: SM00164; TBC; 1.  
 SO SEQUENCE 637 AA; 73289 MW; 08C1A26B6ED37E13 CRC64;

Oy 2 TEVAPT 7  
 |||||  
 Db 181 TEVAPT 186

RESULT 20  
 09WM35 PRELIMINARY; PRT; 652 AA.  
 ID 09WM35  
 AC 09WM35;  
 DT 01-MAY-2000 (TRENBLREL. 13, Created)  
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)  
 DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)  
 DE CG8743 PROTEIN.  
 GN CG8743.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burlis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegan C.,  
 RA Jasthi M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Paclab J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheelers F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhou X., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL, AE003516; AAF9118.1; -;  
 DR FlyBase: FBgn0036904; CG8743.  
 DR InterPro: IPR001682; -;  
 DR InterPro: IPR002111; -;  
 SO SEQUENCE 652 AA; 74251 MW; 21989FCE490235E4 CRC64;

Query Match 50.0%; Score 6; DB 3; Length 637;  
 Best Local Similarity 100.0%; Pred. No. 84;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 50.0%; Score 6; DB 5; Length 652;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE-94150718; PubMed-7906398;  
 RA Wilson R., Alnesough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sprat J., Wohlman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans";  
 RL Nature 368:32-38(1994).  
 DR EMBL; Z99942; CAB17068.1; -;  
 DR InterPro; IPR002035; -;  
 DR Pfam; PF00092; vwa; 1.  
 SO SEQUENCE 460 AA; 49872 MW; 54FA0FAC48B8ADBA CRC64;

Query Match 50.0%; Score 6; DB 5; Length 460;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7  
 |||||  
 DB 73 TPVAPT 78

RESULT 15  
 Q26617 ID Q26617 PRELIMINARY; PRT; 486 AA.  
 AC Q26617;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE SPCGF.  
 OS Strongylocentrotus purpuratus (Purple sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidea; Euechinoidea; Echinacea; Echinoidea; Strongylocentrotidae;  
 OC Strongylocentrotus.  
 NC NCB1\_TaxID=7668;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-95301110; PubMed-7781910;  
 RA Zeller R.W., Coffman J.A., Harrington M.G., Britten R.J.,  
 RA Davidson E.H.;  
 RT "SPGCF1, a sea urchin embryo DNA-binding protein, exists as five  
 RT nested variants encoded by a single mRNA."  
 RL Dev. Biol. 169:713-727(1995).  
 DR EMBL; U18784; AAA86479.1; -;  
 DR InterPro; IPR000255; -;  
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN\_1.  
 SO SEQUENCE 486 AA; 54616 MW; A9ABP23F3F826357 CRC64;

Query Match 50.0%; Score 6; DB 5; Length 486;  
 Best Local Similarity 100.0%; Pred. No. 68;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PTOEVK 11  
 |||||  
 DB 40 PTOEVK 45

RESULT 16  
 Q9DEV2 ID Q9DEV2 PRELIMINARY; PRT; 530 AA.  
 AC Q9DEV2;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE TARGET OF JUN 3.  
 CN TOJ3.  
 OS Coturnix coturnix (common quail).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Coturnix.  
 NC NCB1\_TaxID=9091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bader A.G., Hartl M., Bister K.;  
 RT "TOJ3, a novel potential target of the oncogenic transcription factor  
 Jun."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY007310; AAC16624.1; -;  
 SO SEQUENCE 530 AA; 57929 MW; E460DB89D9FF490B CRC64;

Query Match 50.0%; Score 6; DB 13; Length 530;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVAP 6  
 |||||  
 DB 170 STPVAP 175

RESULT 17  
 Q9UNJ0 ID Q9UNJ0 PRELIMINARY; PRT; 609 AA.  
 AC Q9UNJ0;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE POLYCISTIN-2-LIKE PROTEIN.  
 GN PKD2L2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NC NCB1\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-TESTIS;  
 RA Veldhuisen B., Spruit L., Dauwerse H.G., Breuning M.H., Peters D.J.M.;  
 RT "Genes homologous to the autosomal dominant polycystic kidney disease  
 RT 2 (PKD2) gene," to the EMBL/GenBank/DBJ databases.  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF118125; AAD46478.1; -;  
 DR InterPro; IPR000636; -;  
 DR InterPro; IPR001682; -;  
 DR InterPro; IPR001696; -;  
 DR InterPro; IPR001968; -;  
 DR InterPro; IPR002111; -;  
 DR Pfam; PF00520; Ion\_trans; 1.  
 DR PRINTS; PR00170; NACHANNEL.  
 DR PRODOM; PD003549; -; 1.  
 SO SEQUENCE 609 AA; 71906 MW; D5DC9829B9A41AB CRC64;

Query Match 50.0%; Score 6; DB 4; Length 609;  
 Best Local Similarity 100.0%; Pred. No. 81;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12  
 |||||  
 DB 298 TOEVKK 303

RESULT 18  
 Q9NZM6 ID Q9NZM6 PRELIMINARY; PRT; 624 AA.  
 AC Q9NZM6;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)

OC Actinomycetales; Streptomycineae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA MEDLINE=97000351; PubMed=8843436;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 CC -1. SIMILARITY: BELONGS TO THE ARAC/YLS FAMILY OF TRANSCRIPTIONAL REGULATORS.  
 CC EMBL: AL356592; CAB92194.1; -.  
 DR InterPro: IPR000005; -.  
 DR Pfam: PF00165; HTH\_ARAC.1.  
 DR PRINTS: PR00032; HTHARAC.  
 DR PROSITE: PS01124; HTH\_ARAC\_FAMILY\_2; 1.  
 DR SMART: SM00342; HTH\_ARAC; 1.  
 DR DNA-binding; Transcription regulation.  
 KW SEQUENCE 289 AA; 31511 MW; F709F9CF93C73F56 CRC64;  
 SQ

Query Match 50.0%; Score 6; DB 2; Length 289;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TRVAP 7  
 DB 37 TRVAP 42

RESULT 12  
 O9RM05 PRELIMINARY; PRT; 293 AA.  
 AC O9RM05;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE HYPOTHEICAL 30.8 KDA PROTEIN.  
 GN DR0864.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RI;  
 RA MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zaleski C.,  
 RA Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.";  
 RL Science 286:1571-1577(1999).  
 DR EMBL: AF001940; AAF10447.1; -.  
 DR TIGR: DR0864; -.  
 KW Hypothetical protein.  
 SO SEQUENCE 293 AA; 30849 MW; 4FA7BFEFF12A0E83 CRC64;  
 RL

Query Match 50.0%; Score 6; DB 2; Length 293;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TRVAP 6  
 DB 227 TRVAP 232

RESULT 13  
 O86673 PRELIMINARY; PRT; 447 AA.  
 AC O86673;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DE PUTATIVE LIPOPROTEIN.  
 GN SCA42.17C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Oliver K., Harris D.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA MEDLINE=97000351; PubMed=8843436;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL031182; CAZ0169.1; -.  
 KW Lipoprotein.  
 SQ SEQUENCE 447 AA; 46712 MW; 809E0091B7834D80 CRC64;  
 SQ

Query Match 50.0%; Score 6; DB 2; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TRVAP 7  
 DB 63 TRVAP 68

RESULT 14  
 O9XT09 PRELIMINARY; PRT; 460 AA.  
 AC O9XT09;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 GN H13N06.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Lennard N.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RL

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=96-62333;  
 RA Beldi F.J., Barlow K.L., Murphy G., Parry J.V., Clewley J.P.;  
 RT "A dual subtype B/E HIV-1 infection with a novel V3 loop crown motif  
 among infections acquired in South East Asia and imported into  
 England.";  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO GAG GENE PROTEIN P24 (CORE NUCLEOCAPSID PROTEIN).  
 DR EMBL: AJ224197; CA11887.1; -  
 DR InterPro: IPR000721; -  
 DR Pfam: PF00607; gag\_p24; 1.  
 KM Core protein; Polyprotein.  
 FT NON\_TER 1 1  
 FT NON\_TER 174 174  
 SQ SEQUENCE 174 AA; 19667 MW; D6F18R245B0707DF CRC64;

Query Match  
 Best Local Similarity 50.0%; Score 6; DB 14; Length 174;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12  
 |||||  
 Db 154 TOEVKK 159

RESULT 8  
 O9FU39 PRELIMINARY; PRT; 237 AA.  
 AC O9FU39;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE P0001B06.23 PROTEIN.  
 GN P0001B06.23.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;  
 OC Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 clone: P0001B06.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF002537; BAB16870.1; -  
 SQ SEQUENCE 237 AA; 25810 MW; 972E4F6C57C2E39A CRC64;

Query Match  
 Best Local Similarity 50.0%; Score 6; DB 10; Length 237;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7  
 |||||  
 Db 220 TPVAPT 225

RESULT 9  
 O9KVZ4 PRELIMINARY; PRT; 274 AA.  
 AC O9KVZ4;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE BETA-LACTAMASE PRECURSOR.  
 GN OXA-23.  
 OS Acinetobacter baumannii.

OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;  
 OC Acinetobacter.  
 OX NCBI\_TaxID=470;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RYC 52763/97;  
 RX MEDLINE=20277880; PubMed=10817708;  
 RA Bou G., Oliver A., Martinez-Beltran J.;  
 RT "OXA-24, a novel class D beta-lactamase with carbapenemase activity in  
 an Acinetobacter baumannii clinical strain.";  
 RL Antimicrob. Agents Chemother. 44:1556-1561(2000).  
 DR EMBL: AJ239129; CAB92323.1; -  
 KM Signal.  
 FT SIGNAL 1 7  
 FT SIGNAL POTENTIAL.  
 SQ SEQUENCE 274 AA; 30810 MW; 5343DC532A8AB4E6 CRC64;

Query Match  
 Best Local Similarity 50.0%; Score 6; DB 2; Length 274;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12  
 |||||  
 Db 197 TOEVKK 202

RESULT 10  
 O84512 PRELIMINARY; PRT; 288 AA.  
 AC O84512;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE HYPOTHETICAL 32.2 KDA PROTEIN.  
 GN CT304.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UM-3/CX;  
 RX MEDLINE=95000809; PubMed=9784136;  
 RA Stephens R.S., Kaiman S., Lammell C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 Chlamydia trachomatis.";  
 RL Science 282:754-759(1998).  
 DR EMBL: AE001323; AAC68105.1; -  
 SQ SEQUENCE 288 AA; 32191 MW; 9B5CAE26D04028B5 CRC64;

Query Match  
 Best Local Similarity 50.0%; Score 6; DB 2; Length 288;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVAP 6  
 |||||  
 Db 131 STPVAP 136

RESULT 11  
 O9KYN4 PRELIMINARY; PRT; 289 AA.  
 AC O9KYN4;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE PUTATIVE ARAC-FAMILY TRANSCRIPTIONAL REGULATOR.  
 GN SC9H11.05.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

Query Match 58.3%; Score 7; DB 14; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 VAPTOEV 10  
1111111  
DB 39 VAPTOEV 45

RESULT 4  
O9PM45 PRELIMINARY; PRT; 227 AA.  
AC O9PM45;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE PUTATIVE TONB TRANSPORT PROTEIN.  
GN TONB2 OR C1630.  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
OC Campylobacter.  
OX NCBI\_TaxID=197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCYC 1168;  
RX MEDLINE=20150912; PubMed=10688204;  
RA Parkhill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,  
BA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,  
RA Jorgensen K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,  
RA Whitehead S., Barrett B.G.;  
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
RT reveals hypervariable sequences.";  
RL Nature 403:665-668(2000).  
DR EMBL; AL139079; CAB73618.1; -  
DR InterPro; IPR003538; -  
DR PRINTS; PR01374; TONBPROTEIN.  
SQ SEQUENCE 227 AA; 26298 MW; 7AEFFB148480DCD CRC64;

Query Match 58.3%; Score 7; DB 2; Length 227;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APTOEVK 11  
1111111  
DB 105 APTOEVK 111

RESULT 5  
O9HYT7 PRELIMINARY; PRT; 102 AA.  
AC O9HYT7;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL PROTEIN PA3307.  
GN PA3307.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Steyer C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
DR EMBL; AE004753; AAG06695.1; -  
KM Hypothetical protein.  
SQ SEQUENCE 102 AA; 11114 MW; D92FEE48B895286B CRC64;

Query Match 50.0%; Score 6; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPRQ 8  
1111111  
DB 27 PVAPRQ 32

RESULT 6  
O9X9T5 PRELIMINARY; PRT; 116 AA.  
ID O9X9T5  
AC O9X9T5;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)  
DE PUTATIVE REGULATORY PROTEIN.  
GN SC011.16.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Seeger K., Harris D.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA James K.D., Parkhill J., Barrett B.G., Rajandream M.A.;  
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,  
RA Kinasht H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL; AL096823; CAB46971.1; -  
SQ SEQUENCE 116 AA; 12228 MW; 4B6966EB74C7E6CD CRC64;

Query Match 50.0%; Score 6; DB 2; Length 116;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12  
1111111  
DB 29 TOEVKK 34

RESULT 7  
O93041 PRELIMINARY; PRT; 174 AA.  
ID O93041  
AC O93041;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE GAG POLYPROTEIN [CONTAINS: CORE PROTEIN(S) P24] (FRAGMENT).  
GN GAG.  
OS Human immunodeficiency virus type 1.



```
669 5 41.7 609 14 Q82672 influenza c
970 5 41.7 609 14 Q82673 influenza c
971 5 41.7 609 14 Q82674 influenza c
972 5 41.7 609 14 Q82675 influenza c
973 5 41.7 620 5 Q9Y0D4
974 5 41.7 620 5 Q9NEG8
975 5 41.7 624 4 Q9P2U4
976 5 41.7 626 4 Q9P2U3
977 5 41.7 627 4 Q9H172
978 5 41.7 633 2 Q59526
979 5 41.7 637 2 P72813
980 5 41.7 637 2 Q9W0E6
981 5 41.7 641 5 Q01667
982 5 41.7 641 5 Q9VU01
983 5 41.7 641 14 Q91Q72
984 5 41.7 642 2 Q9FC03
985 5 41.7 645 5 Q9NPF7
986 5 41.7 645 14 Q91Q73
987 5 41.7 645 14 Q91Q69
988 5 41.7 646 2 Q07625
989 5 41.7 646 14 Q67417
990 5 41.7 647 14 Q67386
991 5 41.7 648 14 Q91Q66
992 5 41.7 649 14 Q39661
993 5 41.7 649 14 Q67416
994 5 41.7 649 14 Q67418
995 5 41.7 649 14 Q67419
996 5 41.7 649 14 Q91Q71
997 5 41.7 649 14 Q91Q70
998 5 41.7 650 14 Q91Q67
999 5 41.7 651 4 Q9UF25
1000 5 41.7 651 5 Q18181
```

## ALIGNMENTS

```
RESULT 1
Q9NRU3 PRELIMINARY: PRT: 86 AA.
AC Q9NRU3:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE P60 PROTEIN (FRAGMENT).
GN IAP.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90256283; PubMed=2111287;
RA Kohler S., Leimister-Wachter M., Chakraborty T., Lottspeich F.,
RA Goebel W.;
RT "The gene coding for protein p60 of Listeria monocytogenes and its use
RT as a specific probe for Listeria monocytogenes.";
RL Infect. Immun. 58:1943-1950(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Park S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF179003; AAD55089.1; -.
FT NON_TER 1 1
FT NON_TER 86 86
SQ SEQUENCE 86 AA; 9194 MW; 71F649A817D697F6 CRC64;
```

```
Query Match 100.0%; Score 12; DB 2; Length 86;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 STPVAPTOEVKK 12

```
Db 30 STPVAPTOEVKK 41
|||||
```

```
RESULT 2
Q03493 PRELIMINARY: PRT: 478 AA.
AC Q03493:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PROTEIN P60 PRECURSOR (INVASION-ASSOCIATED PROTEIN).
GN IAP.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93094153; PubMed=1459966;
RA Buberl A., Kuhn M., Goebel W., Koehler S.;
RT "Structural and functional properties of the p60 proteins from
RT different Listeria species.";
RL J. Bacteriol. 174:8166-8171(1992).
CC -!- FUNCTION: THIS MAJOR EXTRACELLULAR PROTEIN MAY BE INVOLVED IN THE
CC INVASION OF NONPROFESSIONAL PHAGOCYTTIC CELLS BY LISTERIA.
DR EMBL; M80351; AAA25280.1; -.
DR InterPro; IPR000064; -.
DR InterPro; IPR002482; -.
DR Pfam; PF00877; NRPC_P60; 1.
DR Pfam; PF01476; LysM; 2.
DR SMART; SM00257; LysM; 1.
KW Signal.
FT SIGNAL 1 27
FT CHAIN 28 478
FT CHAIN 478 478
SQ SEQUENCE 478 AA; 49997 MW; D46D050507402344 CRC64;
```

Query Match 100.0%; Score 12; DB 2; Length 478;

Best Local Similarity 100.0%; Pred. No. 2.7e-05; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 STPVAPTOEVKK 12
Db 146 STPVAPTOEVKK 157
|||||
```

```
RESULT 3
Q67673 PRELIMINARY: PRT: 94 AA.
AC Q67673:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HEXON PROTEIN (FRAGMENT).
OS porcine adenovirus sp.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=47901;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Pring-Akerblom P.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X95630; CAAG4884.1; -.
DR HSSP; P03277; IDHX.
DR InterPro; IPR000736; -.
DR Pfam; PF01065; Adeno_hexon; 1.
KW Hexon protein.
FT NON_TER 1 1
FT NON_TER 94 94
SQ SEQUENCE 94 AA; 10807 MW; 2646FB97DDCC9FD CRC64;
```

823	5	41.7	492	14	Q9WF75	Q9WF75 human immun	896	5	41.7	509	14	056939	056939 human papill
824	5	41.7	492	14	Q9WF72	Q9WF72 human immun	897	5	41.7	509	14	056945	056945 human papill
825	5	41.7	492	14	Q9WF69	Q9WF69 human immun	898	5	41.7	511	2	052569	052569 amycolatops
826	5	41.7	492	14	Q9WF66	Q9WF66 human immun	899	5	41.7	512	5	076725	076725 caenorhabdi
827	5	41.7	492	14	Q9QRW9	Q9QRW9 human immun	900	5	41.7	512	14	091079	091079 human immun
828	5	41.7	492	14	Q9EAL3	Q9EAL3 human immun	901	5	41.7	516	10	Q9SVB0	Q9SVB0 arabisdopsi
829	5	41.7	493	10	024326	024326 phaseolus v	902	5	41.7	516	14	012670	012670 colobus mon
830	5	41.7	493	10	Q9FNR2	Q9FNR2 arabisdopsi	903	5	41.7	516	14	P88804	P88804 human papil
831	5	41.7	493	10	Q9FHL9	Q9FHL9 arabisdopsi	904	5	41.7	517	14	071261	071261 human immun
832	5	41.7	493	14	Q9OMW9	Q9OMW9 human immun	905	5	41.7	517	14	073329	073329 human immun
833	5	41.7	493	14	Q9IV95	Q9IV95 human immun	906	5	41.7	519	5	Q9NEU8	Q9NEU8 caenorhabdi
834	5	41.7	494	14	P90256	P90256 human immun	907	5	41.7	520	11	035731	035731 mus musculu
835	5	41.7	494	14	P90257	P90257 human immun	908	5	41.7	521	2	Q9KJ00	Q9KJ00 arthrobacte
836	5	41.7	494	14	P90259	P90259 human immun	909	5	41.7	526	4	Q9P280	Q9P280 homo sapien
837	5	41.7	494	14	Q9OS00	Q9OS00 human immun	910	5	41.7	528	10	Q9SRE2	Q9SRE2 arabisdopsi
838	5	41.7	494	14	Q9OC00	Q9OC00 human immun	911	5	41.7	530	2	Q9I008	Q9I008 pseudomonas
839	5	41.7	494	14	Q9OB14	Q9OB14 human immun	912	5	41.7	534	1	Q9E459	Q9E459 methanobact
840	5	41.7	494	14	Q9IW44	Q9IW44 human immun	913	5	41.7	534	2	Q9RJ53	Q9RJ53 streptomyc
841	5	41.7	495	14	Q9OP03	Q9OP03 human immun	914	5	41.7	535	14	Q9QWK6	Q9QWK6 chilita viru
842	5	41.7	495	14	Q9OB22	Q9OB22 human immun	915	5	41.7	535	14	Q9IV39	Q9IV39 human calic
843	5	41.7	495	14	Q9IWS3	Q9IWS3 human immun	916	5	41.7	537	2	Q9Z8U8	Q9Z8U8 chlamydia p
844	5	41.7	495	14	Q76126	Q76126 human immun	917	5	41.7	537	2	Q9JRV7	Q9JRV7 chlamydia p
845	5	41.7	496	5	Q17537	Q17537 caenorhabdi	918	5	41.7	539	10	Q80914	Q80914 arabisdopsi
846	5	41.7	496	14	Q9WSG0	Q9WSG0 human immun	919	5	41.7	539	11	060664	060664 mus musculu
847	5	41.7	496	14	Q9WS51	Q9WS51 human immun	920	5	41.7	540	4	060230	060230 homo sapien
848	5	41.7	496	14	Q9WLH9	Q9WLH9 human immun	921	5	41.7	542	5	Q23446	Q23446 caenorhabdi
849	5	41.7	496	14	Q9QST1	Q9QST1 human immun	922	5	41.7	543	10	Q22188	Q22188 arabisdopsi
850	5	41.7	496	14	Q9DKK0	Q9DKK0 human immun	923	5	41.7	544	11	060764	060764 mus musculu
851	5	41.7	497	14	Q41792	Q41792 human immun	924	5	41.7	547	13	Q91349	Q91349 xenopus lae
852	5	41.7	497	14	Q89926	Q89926 human immun	925	5	41.7	548	2	Q9ZEX1	Q9ZEX1 listeria mo
853	5	41.7	497	14	Q9WLL2	Q9WLL2 human immun	926	5	41.7	548	2	Q9EXF6	Q9EXF6 listeria mo
854	5	41.7	497	14	Q9WLL1	Q9WLL1 human immun	927	5	41.7	548	5	Q9XVY9	Q9XVY9 bombyx mori
855	5	41.7	497	14	Q9ID96	Q9ID96 human immun	928	5	41.7	549	5	Q24441	Q24441 drosophila
856	5	41.7	498	14	Q9H7A2	Q9H7A2 homo sapien	929	5	41.7	550	2	Q9SAM6	Q9SAM6 escherichia
857	5	41.7	498	14	P88150	P88150 human immun	930	5	41.7	551	2	Q9KZS2	Q9KZS2 streptomyc
858	5	41.7	498	14	Q77372	Q77372 human immun	931	5	41.7	552	5	Q9V152	Q9V152 drosophila
859	5	41.7	498	14	Q79665	Q79665 human immun	932	5	41.7	553	2	Q9RSN4	Q9RSN4 streptomyc
860	5	41.7	498	14	Q9WIR6	Q9WIR6 human immun	933	5	41.7	555	2	Q9SG12	Q9SG12 deinococcus
861	5	41.7	498	14	P90239	P90239 human immun	934	5	41.7	555	2	Q9R3X4	Q9R3X4 escherichia
862	5	41.7	498	14	Q9QWL1	Q9QWL1 human immun	935	5	41.7	556	2	Q9SAM3	Q9SAM3 escherichia
863	5	41.7	498	14	Q9QWL3	Q9QWL3 human immun	936	5	41.7	556	10	Q9ML72	Q9ML72 arabisdopsi
864	5	41.7	498	14	Q9Q650	Q9Q650 human immun	937	5	41.7	556	10	Q9LPW2	Q9LPW2 arabisdopsi
865	5	41.7	498	14	Q9Q6H6	Q9Q6H6 human immun	938	5	41.7	557	2	Q9R869	Q9R869 aquilex aeo
866	5	41.7	498	14	Q9Q6G9	Q9Q6G9 human immun	939	5	41.7	557	2	Q9K9Z6	Q9K9Z6 bacillus ha
867	5	41.7	498	14	Q9DKH3	Q9DKH3 human immun	940	5	41.7	558	2	P73902	P73902 synechocyst
868	5	41.7	498	14	Q9DH06	Q9DH06 human immun	941	5	41.7	560	2	Q9SAM5	Q9SAM5 escherichia
869	5	41.7	499	11	Q88714	Q88714 mus musculu	942	5	41.7	561	4	Q9Z952	Q9Z952 homo sapien
870	5	41.7	499	14	Q72860	Q72860 human immun	943	5	41.7	562	1	Q9HRL9	Q9HRL9 halobacteri
871	5	41.7	499	14	Q74834	Q74834 human immun	944	5	41.7	567	2	Q93932	Q93932 listeria mo
872	5	41.7	499	14	P89692	P89692 human immun	945	5	41.7	570	2	Q66690	Q66690 aquilex aeo
873	5	41.7	499	14	Q90712	Q90712 human immun	946	5	41.7	572	5	Q26042	Q26042 penaeus jap
874	5	41.7	499	14	Q9IWE6	Q9IWE6 human immun	947	5	41.7	573	3	Q08179	Q08179 saccharomyc
875	5	41.7	500	14	P73537	P73537 synechocyst	948	5	41.7	573	5	Q19963	Q19963 caenorhabdi
876	5	41.7	500	14	P89971	P89971 human immun	949	5	41.7	574	2	Q19963	Q19963 caenorhabdi
877	5	41.7	500	14	Q74806	Q74806 human immun	950	5	41.7	578	5	Q9TVL0	Q9TVL0 synechocyst
878	5	41.7	500	14	Q9YX54	Q9YX54 human immun	951	5	41.7	580	10	Q9ZVU7	Q9ZVU7 arabisdopsi
879	5	41.7	500	14	Q9MO10	Q9MO10 human immun	952	5	41.7	581	2	Q9RU45	Q9RU45 deinococcus
880	5	41.7	502	14	Q9Q6X6	Q9Q6X6 human immun	953	5	41.7	583	2	Q9ZNT9	Q9ZNT9 pseudomonas
881	5	41.7	502	14	Q9IDA5	Q9IDA5 human immun	954	5	41.7	583	5	Q17977	Q17977 caenorhabdi
882	5	41.7	502	14	Q9IC18	Q9IC18 chimpanzee	955	5	41.7	584	4	Q9UNR7	Q9UNR7 homo sapien
883	5	41.7	503	10	Q9LD28	Q9LD28 arabisdopsi	956	5	41.7	585	2	P71246	P71246 escherichia
884	5	41.7	503	11	Q9ROM4	Q9ROM4 mus musculu	957	5	41.7	585	2	Q9L736	Q9L736 escherichia
885	5	41.7	503	11	Q9ESZ1	Q9ESZ1 mus musculu	958	5	41.7	585	2	Q9K2Y6	Q9K2Y6 escherichia
886	5	41.7	503	14	Q97058	Q97058 human immun	959	5	41.7	587	10	Q9MOT2	Q9MOT2 arabisdopsi
887	5	41.7	504	2	Q9ZBWS	Q9ZBWS streptomyc	960	5	41.7	596	10	Q9LSW2	Q9LSW2 arabisdopsi
888	5	41.7	505	2	Q9ZS08	Q9ZS08 helicobacte	961	5	41.7	599	2	Q9PP00	Q9PP00 campylobact
889	5	41.7	505	4	Q14395	Q14395 homo sapien	962	5	41.7	604	2	Q00509	Q00509 streptomyc
890	5	41.7	505	4	Q9UNR8	Q9UNR8 homo sapien	963	5	41.7	605	1	Q26511	Q26511 methanobact
891	5	41.7	507	4	Q13064	Q13064 homo sapien	964	5	41.7	606	2	Q53561	Q53561 streptococc
892	5	41.7	507	14	Q9Q079	Q9Q079 chimpanzee	965	5	41.7	609	13	Q13027	Q13027 xenopus lae
893	5	41.7	508	5	Q93227	Q93227 caenorhabdi	966	5	41.7	609	14	Q82659	Q82659 influenza c
894	5	41.7	508	6	Q9GMC7	Q9GMC7 felis silve	967	5	41.7	609	14	Q82670	Q82670 influenza c
895	5	41.7	509	6	Q9GMC7	Q9GMC7 bison bison	968	5	41.7	609	14	Q82671	Q82671 influenza c

677	5	41.7	385	6	09GL14	09g114 microptamo	750	5	41.7	469	2	09ZL71	09z171 helicobacte
678	5	41.7	385	10	09%P00	09%P00 brassica ju	751	5	41.7	469	14	09JXN9	09jxn9 human immun
679	5	41.7	385	14	080627	080627 human immun	752	5	41.7	469	14	09END0	09end0 human immun
680	5	41.7	386	5	P91632	P91632 drosophila	753	5	41.7	469	14	09ENC2	09enc2 human immun
681	5	41.7	387	4	09H8P7	09h8p7 homo sapien	754	5	41.7	469	14	09ENC0	09enc0 human immun
682	5	41.7	388	5	09VUD1	09vud1 drosophila	755	5	41.7	469	14	09ENC0	09enc0 human immun
683	5	41.7	389	2	09K7L3	09k7l3 bacillus ha	756	5	41.7	470	5	09ENB8	09enb8 human immun
684	5	41.7	392	4	09NPL5	09npl5 homo sapien	757	5	41.7	470	5	09ENB4	09enb4 human immun
685	5	41.7	392	5	09VW77	09vw77 drosophila	758	5	41.7	471	5	09V8B0	09v8b0 drosophila
686	5	41.7	393	14	041549	041549 human immun	759	5	41.7	472	14	09ENB4	09enb4 human immun
687	5	41.7	394	5	09VMU4	09vmu4 drosophila	760	5	41.7	472	14	09ENB9	09enb9 human immun
688	5	41.7	394	5	09XU00	09xu00 caenorhabd1	761	5	41.7	473	2	050210	050210 synechococ
689	5	41.7	395	2	09KE01	09ke01 bacillus ha	762	5	41.7	473	14	09ENB0	09enb0 human immun
690	5	41.7	398	4	09UGL3	09ugl3 homo sapien	763	5	41.7	473	14	09ENAB	09enab human immun
691	5	41.7	400	11	09EQU2	09equ2 apodemus ag	764	5	41.7	473	14	09ENAB	09enab human immun
692	5	41.7	406	3	09UYD9	09uyd9 alternaria	765	5	41.7	473	14	09EN97	09en97 human immun
693	5	41.7	407	5	09VBB4	09vbb4 drosophila	766	5	41.7	473	14	09EN94	09en94 human immun
694	5	41.7	408	2	09FD10	09fd10 salmoneila	767	5	41.7	474	14	09EN90	09en90 human immun
695	5	41.7	408	4	09UGL4	09ugl4 homo sapien	768	5	41.7	475	2	09X727	09x727 escherichia
696	5	41.7	409	2	09LAA3	09l443 shigella dy	769	5	41.7	475	2	09F8R9	09f8r9 escherichia
697	5	41.7	410	2	005446	005446 mycobacteri	770	5	41.7	475	2	09F8R3	09f8r3 escherichia
698	5	41.7	410	2	09FAK8	09fak8 rhizobium l	771	5	41.7	475	2	09EY22	09ey22 escherichia
699	5	41.7	410	10	09MA98	09ma98 arabidopsis	772	5	41.7	475	10	09LNM7	09lnw7 arabidopsis
700	5	41.7	411	2	09FAK7	09fak7 rhizobium l	773	5	41.7	476	5	P91292	P91292 caenorhabd1
701	5	41.7	411	5	09V943	09v943 drosophila	774	5	41.7	476	5	003996	003996 plasmodium
702	5	41.7	411	14	081814	081814 hepatitis c	775	5	41.7	476	14	09EN92	09en92 human immun
703	5	41.7	414	14	081329	081329 hepatitis c	776	5	41.7	477	2	09AXT6	09axt6 delnrococcus
704	5	41.7	414	14	P89958	P89958 hepatitis c	777	5	41.7	478	5	025684	025684 plasmodium
705	5	41.7	414	14	P89959	P89959 hepatitis c	778	5	41.7	481	2	09LA08	09laq8 clostridium
706	5	41.7	415	14	081554	081554 hepatitis c	779	5	41.7	481	14	077697	077697 human immun
707	5	41.7	416	2	031444	031444 bacillus su	780	5	41.7	482	5	09LVI9	09lvi9 human immun
708	5	41.7	418	2	09K9F3	09k9f3 bacillus ha	781	5	41.7	482	5	044196	044196 caenorhabd1
709	5	41.7	418	4	075066	075066 homo sapien	782	5	41.7	482	5	09XY38	09xy38 acanthamoeb
710	5	41.7	418	4	09NDA8	09nda8 homo sapien	783	5	41.7	482	5	003998	003998 plasmodium
711	5	41.7	421	2	09Z740	09z740 neisseria m	784	5	41.7	484	5	0909P0	0909p0 hydra atten
712	5	41.7	421	2	09J034	09j034 neisseria m	785	5	41.7	484	14	077699	077699 human immun
713	5	41.7	421	4	09UNN6	09un66 homo sapien	786	5	41.7	484	14	077700	077700 human immun
714	5	41.7	421	4	09UBH9	09ubh9 homo sapien	787	5	41.7	484	14	080622	080622 human immun
715	5	41.7	421	4	09Y6I7	09y6i7 homo sapien	788	5	41.7	485	5	09N4N9	09n4n9 caenorhabd1
716	5	41.7	422	2	P96Z08	P96Z08 mycobacteri	789	5	41.7	486	4	09UCX4	09ucx4 homo sapien
717	5	41.7	423	1	093642	093642 halobacteri	790	5	41.7	486	14	077804	077804 human immun
718	5	41.7	423	1	09HGT5	09hgt5 halobacteri	791	5	41.7	486	14	078566	078566 human immun
719	5	41.7	423	2	049939	049939 mycobacteri	792	5	41.7	486	14	079451	079451 human immun
720	5	41.7	423	5	09UIM3	09uim3 drosophila	793	5	41.7	486	14	080628	080628 human immun
721	5	41.7	429	2	09Z4B1	09z4b1 escherichia	794	5	41.7	486	14	080629	080629 human immun
722	5	41.7	429	2	09R2H0	09r2h0 salmonella	795	5	41.7	486	14	080634	080634 human immun
723	5	41.7	434	2	09X6H2	09x6h2 streptococ	796	5	41.7	487	14	080621	080621 human immun
724	5	41.7	434	14	089933	089933 human immun	797	5	41.7	488	10	09LNX1	09lnx1 arabidopsis
725	5	41.7	435	5	093322	093322 caenorhabd1	798	5	41.7	488	14	079362	079362 human immun
726	5	41.7	436	14	091V79	091v79 human immun	799	5	41.7	488	14	09WT03	09wt03 human herpe
727	5	41.7	438	5	09NG13	09ng13 branchiosto	800	5	41.7	489	14	079350	079350 human immun
728	5	41.7	439	5	09V9N5	09v9n5 drosophila	801	5	41.7	489	14	079354	079354 human immun
729	5	41.7	440	2	025464	025464 helicobacte	802	5	41.7	489	14	079360	079360 human immun
730	5	41.7	443	1	058579	058579 pyrococcus	803	5	41.7	489	14	080306	080306 human immun
731	5	41.7	444	5	026861	026861 trypansom	804	5	41.7	489	14	080307	080307 human immun
732	5	41.7	444	5	09VM63	09vm63 drosophila	805	5	41.7	489	14	080631	080631 human immun
733	5	41.7	445	2	045294	045294 bacteroides	806	5	41.7	489	14	080305	080305 human immun
734	5	41.7	445	2	09KTI9	09kti9 bacteroides	807	5	41.7	490	2	09RMV1	09rmv1 bacillus an
735	5	41.7	446	5	09N3V2	09n3v2 caenorhabd1	808	5	41.7	490	14	080623	080623 human immun
736	5	41.7	448	5	09GMF9	09gmf9 drosophila	809	5	41.7	490	14	P90073	P90073 human immun
737	5	41.7	448	13	09Z019	09z019 xenopus lae	810	5	41.7	490	14	P90255	P90255 human immun
738	5	41.7	452	5	027696	027696 musca domes	811	5	41.7	490	14	09WC83	09wc83 human immun
739	5	41.7	456	11	09JIT8	09jit8 mus musculu	812	5	41.7	491	2	09F5I8	09f5i8 alternomonas
740	5	41.7	459	14	09ZHC0	09zhc0 streptococ	813	5	41.7	491	14	080181	080181 human immun
741	5	41.7	459	14	09JRS3	09jrs3 human papil	814	5	41.7	491	14	09WLH4	09wlh4 human immun
742	5	41.7	461	5	045051	045051 caenorhabd1	815	5	41.7	491	14	09QEF9	09qef9 human immun
743	5	41.7	462	10	09W143	09w143 drosophila	816	5	41.7	491	14	09Q8Z6	09q8z6 human immun
744	5	41.7	462	10	09LPM4	09lpw4 arabidopsis	817	5	41.7	492	2	09ZND0	09znd0 clostridium
745	5	41.7	464	2	P74828	P74828 sphingomona	818	5	41.7	492	14	079351	079351 human immun
746	5	41.7	465	14	090729	090729 human papil	819	5	41.7	492	14	074452	074452 human immun
747	5	41.7	467	14	09NXX1	09nxx1 homo sapien	820	5	41.7	492	14	089291	089291 human immun
748	5	41.7	467	14	040461	040461 human immun	821	5	41.7	492	14	089958	089958 human immun
749	5	41.7	468	14	09ENB2	09enb2 human immun	822	5	41.7	492	14	089962	089962 human immun

531	5	41.7	255	14	Q9J602	Q9J602 human immun	604	5	41.7	315	5	018121	018121 caenorhabdi
532	5	41.7	256	4	Q9X611	Q9X611 streptococc	605	5	41.7	315	5	09XW49	09XW49 caenorhabdi
533	5	41.7	256	4	Q9NUR3	Q9NUR3 homo sapien	606	5	41.7	318	2	066259	066259 actinobacti
534	5	41.7	256	14	Q9J709	Q9J709 human immun	607	5	41.7	318	2	005380	005380 actinobacti
535	5	41.7	256	14	Q9J607	Q9J607 human immun	608	5	41.7	323	11	Q9R292	Q9R292 mus muscu
536	5	41.7	256	14	Q9J607	Q9J607 human immun	609	5	41.7	327	2	09PAP3	09PAP3 xylella fas
537	5	41.7	257	14	Q9J765	Q9J765 human immun	610	5	41.7	327	10	Q9LD07	Q9LD07 oryza sativ
538	5	41.7	257	14	Q9J6W3	Q9J6W3 human immun	611	5	41.7	327	11	Q9OWU6	Q9OWU6 mus muscu
539	5	41.7	257	14	Q9J6P0	Q9J6P0 human immun	612	5	41.7	329	5	Q9Y205	Q9Y205 ephydactia f
540	5	41.7	258	2	Q9K410	Q9K410 streptomyc	613	5	41.7	329	5	Q9Y111	Q9Y111 drosophila
541	5	41.7	258	14	Q9J6R3	Q9J6R3 human immun	614	5	41.7	330	14	Q72628	Q72628 human immun
542	5	41.7	259	14	Q9J6P4	Q9J6P4 human immun	615	5	41.7	335	2	Q9RX81	Q9RX81 delinococcus
543	5	41.7	259	14	Q9J6M3	Q9J6M3 thermoplasm	616	5	41.7	335	2	Q9RSC0	Q9RSC0 delinococcus
544	5	41.7	259	14	Q10736	Q10736 human immun	617	5	41.7	335	5	Q9U4D3	Q9U4D3 caenorhabdi
545	5	41.7	260	5	Q17625	Q17625 caenorhabdi	618	5	41.7	335	5	Q9NEF0	Q9NEF0 caenorhabdi
546	5	41.7	260	7	P79551	P79551 homo sapien	619	5	41.7	335	5	Q9NEF6	Q9NEF6 caenorhabdi
547	5	41.7	261	3	Q9P390	Q9P390 neurospora	620	5	41.7	335	5	Q9N6L3	Q9N6L3 caenorhabdi
548	5	41.7	261	7	Q30096	Q30096 homo sapien	621	5	41.7	337	2	Q53089	Q53089 lactobacilli
549	5	41.7	261	7	Q30099	Q30099 homo sapien	622	5	41.7	339	2	Q9RXE8	Q9RXE8 delinococcus
550	5	41.7	261	7	Q30061	Q30061 homo sapien	623	5	41.7	339	14	Q9IE68	Q9IE68 human immun
551	5	41.7	261	7	Q31633	Q31633 homo sapien	624	5	41.7	342	2	Q925F8	Q925F8 mycobacteri
552	5	41.7	261	14	Q9J6R5	Q9J6R5 human immun	625	5	41.7	343	2	Q9RKG7	Q9RKG7 vibrio chol
553	5	41.7	261	14	Q9J6P5	Q9J6P5 human immun	626	5	41.7	343	8	Q9Z238	Q9Z238 trichophyto
554	5	41.7	261	14	Q9J6N6	Q9J6N6 human immun	627	5	41.7	344	14	Q9IWA5	Q9IWA5 human immun
555	5	41.7	262	14	Q9J736	Q9J736 human immun	628	5	41.7	347	5	Q9VRJ9	Q9VRJ9 drosophila
556	5	41.7	263	10	P94481	P94481 bacillus su	629	5	41.7	351	5	Q9YST8	Q9YST8 drosophila
557	5	41.7	263	10	Q9FLK5	Q9FLK5 arabidopsis	630	5	41.7	351	10	Q23142	Q23142 arabidopsis
558	5	41.7	264	7	Q29826	Q29826 homo sapien	631	5	41.7	353	5	Q9W418	Q9W418 drosophila
559	5	41.7	264	7	Q29970	Q29970 homo sapien	632	5	41.7	353	14	Q9ENC6	Q9ENC6 human immun
560	5	41.7	264	14	Q10730	Q10730 human immun	633	5	41.7	353	14	Q9ENC4	Q9ENC4 human immun
561	5	41.7	265	14	Q10729	Q10729 human immun	634	5	41.7	354	5	Q9NHP6	Q9NHP6 dictyosteli
562	5	41.7	266	5	Q15795	Q15795 homo sapien	635	5	41.7	355	2	Q67685	Q67685 dictyosteli
563	5	41.7	266	5	Q25732	Q25732 plasmodium	636	5	41.7	356	2	Q9S5G6	Q9S5G6 escherichia
564	5	41.7	267	5	Q16704	Q16704 caenorhabdi	637	5	41.7	359	2	Q9MWN3	Q9MWN3 synechococc
565	5	41.7	267	14	Q87048	Q87048 semliki for	638	5	41.7	359	8	Q37749	Q37749 cephalospor
566	5	41.7	268	14	Q10735	Q10735 human immun	639	5	41.7	360	4	Q9UGM6	Q9UGM6 homo sapien
567	5	41.7	269	5	Q30155	Q30155 homo sapien	640	5	41.7	362	2	Q9KKG9	Q9KKG9 yersinia en
568	5	41.7	271	5	Q21382	Q21382 caenorhabdi	641	5	41.7	362	4	Q9UHS5	Q9UHS5 homo sapien
569	5	41.7	273	2	Q05936	Q05936 pseudomonas	642	5	41.7	362	5	Q9NEZ1	Q9NEZ1 caenorhabdi
570	5	41.7	275	4	Q9NRB1	Q9NRB1 homo sapien	643	5	41.7	362	10	Q82054	Q82054 saccharum o
571	5	41.7	276	2	P73137	P73137 synechocyst	644	5	41.7	363	10	Q9M6B4	Q9M6B4 vilis vinif
572	5	41.7	276	5	Q94215	Q94215 caenorhabdi	645	5	41.7	366	14	Q9M6B3	Q9M6B3 reovirus sp
573	5	41.7	278	2	Q68435	Q68435 leptospira	646	5	41.7	366	14	Q86314	Q86314 reovirus sp
574	5	41.7	278	2	Q68437	Q68437 leptospira	647	5	41.7	366	14	Q86315	Q86315 reovirus sp
575	5	41.7	282	5	Q9U221	Q9U221 caenorhabdi	648	5	41.7	366	14	Q86316	Q86316 reovirus sp
576	5	41.7	286	2	Q9R640	Q9R640 mycobacteri	649	5	41.7	366	14	Q86317	Q86317 reovirus sp
577	5	41.7	287	1	Q9V1J3	Q9V1J3 pyrococcus	650	5	41.7	366	14	Q86318	Q86318 reovirus sp
578	5	41.7	288	2	Q9JXF2	Q9JXF2 neisseria m	651	5	41.7	366	14	Q86319	Q86319 reovirus sp
579	5	41.7	289	2	Q67230	Q67230 aquifex aeo	652	5	41.7	366	14	Q86320	Q86320 reovirus sp
580	5	41.7	291	2	Q06449	Q06449 serratia ma	653	5	41.7	366	14	Q86323	Q86323 reovirus sp
581	5	41.7	291	2	Q07464	Q07464 salmonella	654	5	41.7	366	14	Q86307	Q86307 reovirus sp
582	5	41.7	293	2	Q9ZJ08	Q9ZJ08 rhodococcus	655	5	41.7	366	14	Q86308	Q86308 reovirus sp
583	5	41.7	293	2	Q9PDL4	Q9PDL4 xylella fas	656	5	41.7	366	14	Q86309	Q86309 reovirus sp
584	5	41.7	293	4	Q14836	Q14836 homo sapien	657	5	41.7	366	14	Q86310	Q86310 reovirus sp
585	5	41.7	293	10	Q49020	Q49020 gossypium h	658	5	41.7	366	14	Q86311	Q86311 reovirus sp
586	5	41.7	296	10	Q22812	Q22812 arabidopsis	659	5	41.7	366	14	Q86312	Q86312 reovirus sp
587	5	41.7	297	5	Q9V1D4	Q9V1D4 drosophila	660	5	41.7	366	14	Q9ENM1	Q9ENM1 reovirus (c
588	5	41.7	298	2	Q9X9Y6	Q9X9Y6 streptomyc	661	5	41.7	367	14	Q87097	Q87097 chimpanzee
589	5	41.7	298	5	Q9VU90	Q9VU90 drosophila	662	5	41.7	367	14	Q41165	Q41165 parametium
590	5	41.7	299	2	Q9KPI8	Q9KPI8 vibrio chol	663	5	41.7	368	2	Q9LJ44	Q9LJ44 actinobact
591	5	41.7	302	10	Q9SAH7	Q9SAH7 arabidopsis	664	5	41.7	368	2	Q9F010	Q9F010 pseudomonas
592	5	41.7	304	14	Q9J4W8	Q9J4W8 human immun	665	5	41.7	369	5	Q9N405	Q9N405 caenorhabdi
593	5	41.7	305	2	Q9K9W1	Q9K9W1 bacillus ha	666	5	41.7	372	14	Q9PYX0	Q9PYX0 xestia c-nl
594	5	41.7	306	10	Q9SDM3	Q9SDM3 prunus dulc	667	5	41.7	375	10	Q39235	Q39235 arabidopsis
595	5	41.7	306	10	Q9M5Q3	Q9M5Q3 petunia hyb	668	5	41.7	375	10	Q49349	Q49349 arabidopsis
596	5	41.7	307	2	Q55410	Q55410 synechocyst	669	5	41.7	376	5	Q9VB07	Q9VB07 drosophila
597	5	41.7	308	2	Q9RUQ2	Q9RUQ2 delinococcus	670	5	41.7	377	11	Q9Z2K3	Q9Z2K3 cricetus
598	5	41.7	309	5	Q9VUV0	Q9VUV0 drosophila	671	5	41.7	377	11	Q9OXK0	Q9OXK0 rattus norv
599	5	41.7	310	11	Q9EBQ7	Q9EBQ7 mus musculu	672	5	41.7	378	2	Q55307	Q55307 streptococc
600	5	41.7	311	10	Q22805	Q22805 arabidopsis	673	5	41.7	380	1	Q9HR58	Q9HR58 halobacteri
601	5	41.7	312	14	Q9WIR5	Q9WIR5 human immun	674	5	41.7	382	4	Q9H878	Q9H878 homo sapien
602	5	41.7	312	14	Q9WIR2	Q9WIR2 human immun	675	5	41.7	382	5	Q9U3E1	Q9U3E1 caenorhabdi
603	5	41.7	314	10	Q9XFF2	Q9XFF2 pyrus pyrif	676	5	41.7	384	4	Q9S977	Q9S977 homo sapien





93	5	41.7	130	14	09GKW5	09GKW5 human immun	166	5	41.7	185	5	09V916	09V916 drosophila
94	5	41.7	130	14	09GKW1	09GKW1 human immun	167	5	41.7	185	10	09MA15	09MA15 arebidopsis
95	5	41.7	131	14	09QL44	09QL44 human immun	168	5	41.7	186	14	09WK89	09WK89 human immun
96	5	41.7	131	14	09QL18	09QL18 human immun	169	5	41.7	186	14	09J6Z5	09J6Z5 human immun
97	5	41.7	131	14	09QL16	09QL16 human immun	170	5	41.7	187	5	09N344	09N344 caenorhabd
98	5	41.7	131	14	09QL08	09QL08 human immun	171	5	41.7	187	14	09JUP3	09JUP3 human immun
99	5	41.7	131	14	09GKX9	09GKX9 human immun	172	5	41.7	190	1	09VIG6	09VIG6 pyrococcus
100	5	41.7	131	14	09GKX7	09GKX7 human immun	173	5	41.7	190	14	093038	093038 human immun
101	5	41.7	131	14	09GKW8	09GKW8 human immun	174	5	41.7	190	14	09J6N5	09J6N5 human immun
102	5	41.7	131	14	09GKW6	09GKW6 human immun	175	5	41.7	191	2	09XCK6	09XCK6 streptococc
103	5	41.7	132	2	09K2C0	09K2C0 chlamydia p	176	5	41.7	192	14	081390	081390 hepatitis c
104	5	41.7	132	14	09QL23	09QL23 human immun	177	5	41.7	192	14	09J733	09J733 human immun
105	5	41.7	133	2	087645	087645 methylococc	178	5	41.7	192	14	09J6N9	09J6N9 human immun
106	5	41.7	133	3	09UTR5	09UTR5 schizosacch	179	5	41.7	193	14	09J7B3	09J7B3 human immun
107	5	41.7	133	14	09OL05	09OL05 human immun	180	5	41.7	193	14	09L5W8	09L5W8 streptococc
108	5	41.7	133	14	09GKX5	09GKX5 human immun	181	5	41.7	195	14	09J769	09J769 human immun
109	5	41.7	134	2	09EJY9	09EJY9 chlamydia m	182	5	41.7	195	14	09J731	09J731 human immun
110	5	41.7	137	5	09N3U4	09N3U4 caenorhabd	183	5	41.7	195	14	09J6U1	09J6U1 human immun
111	5	41.7	138	14	09EBV4	09EBV4 foot-and-mo	184	5	41.7	195	14	09JUQ4	09JUQ4 human immun
112	5	41.7	138	14	09EBV2	09EBV2 foot-and-mo	185	5	41.7	196	14	09J796	09J796 human immun
113	5	41.7	138	14	09EBT3	09EBT3 foot-and-mo	186	5	41.7	197	2	09RKU2	09RKU2 streptomyc
114	5	41.7	140	5	09NSB1	09NSB1 caenorhabd	187	5	41.7	197	14	09JUN4	09JUN4 human immun
115	5	41.7	141	1	09YAL8	09YAL8 aeropyrum p	188	5	41.7	198	14	09J787	09J787 human immun
116	5	41.7	141	3	005690	005690 saccharomyc	189	5	41.7	198	14	09JUL8	09JUL8 human immun
117	5	41.7	142	14	09J6S4	09J6S4 human immun	190	5	41.7	198	14	09JUL3	09JUL3 human immun
118	5	41.7	143	2	007266	007266 mycobacteri	191	5	41.7	198	14	09JUL2	09JUL2 human immun
119	5	41.7	147	2	09WYK3	09WYK3 thermotoga	192	5	41.7	199	2	09K531	09K531 pseudomonas
120	5	41.7	147	5	076560	076560 caenorhabd	193	5	41.7	199	14	09J6Z7	09J6Z7 human immun
121	5	41.7	148	6	09GKP9	09GKP9 bos taurus	194	5	41.7	199	14	09JUP4	09JUP4 human immun
122	5	41.7	148	10	09FPI3	09FPI3 arebidopsis	195	5	41.7	199	14	09JUM8	09JUM8 human immun
123	5	41.7	149	14	068732	068732 hepatitis c	196	5	41.7	200	14	056247	056247 tobaccco rat
124	5	41.7	152	14	09J756	09J756 human immun	197	5	41.7	200	14	09J6Q2	09J6Q2 human immun
125	5	41.7	156	6	09N146	09N146 macaca mula	198	5	41.7	200	14	09JUP2	09JUP2 human immun
126	5	41.7	157	2	049991	049991 mycobacteri	199	5	41.7	201	3	P87261	P87261 saccharomyc
127	5	41.7	157	5	09GKJ6	09GKJ6 caenorhabd	200	5	41.7	201	14	09J744	09J744 human immun
128	5	41.7	161	2	09FBC1	09FBC1 streptococc	201	5	41.7	201	14	09JVR7	09JVR7 human immun
129	5	41.7	161	10	09SSS3	09SSS3 arebidopsis	202	5	41.7	201	14	09JUS1	09JUS1 human immun
130	5	41.7	161	14	09J771	09J771 human immun	203	5	41.7	201	14	09JUL1	09JUL1 human immun
131	5	41.7	161	14	09J6U4	09J6U4 human immun	204	5	41.7	202	14	09J7A2	09J7A2 human immun
132	5	41.7	162	14	09J6U0	09J6U0 human immun	205	5	41.7	202	14	09JUT1	09JUT1 human immun
133	5	41.7	163	5	09N3V0	09N3V0 caenorhabd	206	5	41.7	202	14	09JUP6	09JUP6 human immun
134	5	41.7	165	2	052819	052819 rhizobium l	207	5	41.7	202	14	09JUM4	09JUM4 human immun
135	5	41.7	165	14	09J788	09J788 human immun	208	5	41.7	202	14	09JUL9	09JUL9 human immun
136	5	41.7	165	14	09J752	09J752 human immun	209	5	41.7	202	14	09JUL0	09JUL0 human immun
137	5	41.7	166	10	09M4T9	09M4T9 homo sapien	210	5	41.7	203	14	09J6T3	09J6T3 human immun
138	5	41.7	166	14	09J7A7	09J7A7 human immun	211	5	41.7	203	14	09JUT0	09JUT0 human immun
139	5	41.7	166	14	09J6V7	09J6V7 human immun	212	5	41.7	203	14	09JUT0	09JUT0 human immun
140	5	41.7	167	14	09J6U3	09J6U3 human immun	213	5	41.7	203	14	09JUT9	09JUT9 human immun
141	5	41.7	170	2	005551	005551 listeria mo	214	5	41.7	203	14	09JUT9	09JUT9 human immun
142	5	41.7	171	5	09N927	09N927 trypanosoma	215	5	41.7	204	14	09Q2J1	09Q2J1 human immun
143	5	41.7	172	14	09J784	09J784 human immun	216	5	41.7	204	14	09J7A8	09J7A8 human immun
144	5	41.7	172	14	09J6S9	09J6S9 human immun	217	5	41.7	204	14	09J797	09J797 human immun
145	5	41.7	173	14	09J781	09J781 human immun	218	5	41.7	204	14	09J782	09J782 human immun
146	5	41.7	173	14	09J6Z3	09J6Z3 human immun	219	5	41.7	204	14	09J741	09J741 human immun
147	5	41.7	175	5	09V9R6	09V9R6 drosophila	220	5	41.7	204	14	09J6Q4	09J6Q4 human immun
148	5	41.7	175	14	09XMF4	09XMF4 sorghum bic	221	5	41.7	204	14	09J6U3	09J6U3 human immun
149	5	41.7	175	14	09J6M9	09J6M9 human immun	222	5	41.7	204	14	09JUT4	09JUT4 human immun
150	5	41.7	177	10	043478	043478 hordeum vul	223	5	41.7	205	14	09J6P1	09J6P1 human immun
151	5	41.7	177	14	09J791	09J791 human immun	224	5	41.7	205	14	09JUT5	09JUT5 human immun
152	5	41.7	178	14	09J751	09J751 human immun	225	5	41.7	205	14	09JUT5	09JUT5 human immun
153	5	41.7	178	14	09J751	09J751 human immun	226	5	41.7	205	14	09JUT0	09JUT0 human immun
154	5	41.7	178	14	09J750	09J750 human immun	227	5	41.7	205	14	09JUT6	09JUT6 human immun
155	5	41.7	178	14	09J747	09J747 human immun	228	5	41.7	205	14	09JUT0	09JUT0 human immun
156	5	41.7	179	10	09J6D4	09J6D4 arebidopsis	229	5	41.7	205	14	09JUT2	09JUT2 human immun
157	5	41.7	179	14	09J786	09J786 human immun	230	5	41.7	205	14	09JUT0	09JUT0 human immun
158	5	41.7	179	14	09J6V2	09J6V2 human immun	231	5	41.7	206	2	09J6S1	09J6S1 pseudomonas
159	5	41.7	180	10	09W9M0	09W9M0 arebidopsis	232	5	41.7	206	2	09EX31	09EX31 streptomyc
160	5	41.7	180	14	09J737	09J737 human immun	233	5	41.7	206	14	09J772	09J772 human immun
161	5	41.7	181	14	09J7A3	09J7A3 human immun	234	5	41.7	206	14	09J6W1	09J6W1 human immun
162	5	41.7	181	14	09J793	09J793 human immun	235	5	41.7	206	14	09J6V6	09J6V6 human immun
163	5	41.7	181	14	09J780	09J780 human immun	236	5	41.7	206	14	09JUT3	09JUT3 human immun
164	5	41.7	183	14	09J706	09J706 human immun	237	5	41.7	207	14	09W9W6	09W9W6 human immun
165	5	41.7	184	14	09J6T7	09J6T7 human immun	238	5	41.7	207	14	09W9W4	09W9W4 human immun

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:42:45 ; Search time 58.01 Seconds  
(without alignments)  
27.369 Million cell updates/sec

Title: US-09-372-036-30  
Perfect score: 12  
Sequence: 1 STEVAPTOQEVKK 12

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL\_16:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mmc:\*
- 8: sp\_organelle:\*
- 9: sp\_plant:\*
- 10: sp\_protist:\*
- 11: sp\_rhodent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	100.0	86	2	Q9RNU3
2	12	100.0	478	2	Q03493
3	7	58.3	94	14	Q67673
4	7	58.3	227	2	Q9PM45
5	6	50.0	102	2	Q9HYT7
6	6	50.0	116	2	Q9X9T5
7	6	50.0	174	14	Q93041
8	6	50.0	237	10	Q9FU39
9	6	50.0	274	2	Q9KVZ4
10	6	50.0	288	2	Q84512
11	6	50.0	289	2	Q9RYN4
12	6	50.0	293	2	Q9RM05
13	6	50.0	447	2	Q86673
14	6	50.0	460	5	Q9YTO9
15	6	50.0	486	5	Q26617
16	6	50.0	530	13	Q9DEV2
17	6	50.0	609	4	Q9UNJ0
18	6	50.0	624	4	Q9NZM6
19	6	50.0	637	3	Q08484

20	50.0	652	5	Q9WV35	Q9WV35 drosophila
21	50.0	748	10	Q22907	Q22907 arabidopsis
22	50.0	750	3	Q9HFZ4	Q9HFZ4 candida alb
23	50.0	775	5	Q9VB49	Q9VB49 drosophila
24	50.0	792	5	Q9UAG1	Q9UAG1 epiphydactia f
25	50.0	819	5	Q9VVA8	Q9VVA8 drosophila
26	50.0	904	5	Q9VEX7	Q9VEX7 drosophila
27	50.0	912	5	Q9NHCL	Q9NHCL drosophila
28	50.0	912	5	Q9NH99	Q9NH99 drosophila
29	50.0	916	5	Q9NHCC	Q9NHCC drosophila
30	50.0	916	5	Q9NHCO	Q9NHCO drosophila
31	50.0	988	14	Q98830	Q98830 human foamy
32	50.0	988	14	Q87041	Q87041 simian foam
33	50.0	989	14	P90288	P90288 human spuma
34	50.0	1052	5	Q9YS50	Q9YS50 drosophila
35	50.0	1230	2	Q25772	Q25772 helicobacte
36	50.0	1633	2	Q9LBB3	Q9LBB3 streptococc
37	50.0	2037	13	Q9PVZ2	Q9PVZ2 xenopus lae
38	41.7	23	11	Q63334	Q63334 rattus norv
39	41.7	52	6	Q9GL38	Q9GL38 bos taurus
40	41.7	62	2	Q47006	Q47006 escherichia
41	41.7	75	2	Q9RRH1	Q9RRH1 delnococtus
42	41.7	77	14	Q91692	Q91692 human immun
43	41.7	82	5	Q17728	Q17728 caenorhabdi
44	41.7	88	5	P91335	P91335 caenorhabdi
45	41.7	88	14	Q9DR28	Q9DR28 human immun
46	41.7	91	14	Q9WJ74	Q9WJ74 human immun
47	41.7	95	14	Q68302	Q68302 hepatitis c
48	41.7	102	14	Q68301	Q68301 hepatitis c
49	41.7	108	14	Q91E25	Q91E25 human immun
50	41.7	112	4	Q9UG25	Q9UG25 homo sapien
51	41.7	113	14	Q93051	Q93051 human immun
52	41.7	113	14	Q93052	Q93052 human immun
53	41.7	113	14	Q93053	Q93053 human immun
54	41.7	113	14	Q9WKA3	Q9WKA3 human immun
55	41.7	113	14	Q9WKA2	Q9WKA2 human immun
56	41.7	113	14	Q9WKA1	Q9WKA1 human immun
57	41.7	113	14	Q9WKA8	Q9WKA8 human immun
58	41.7	113	14	Q9WKA7	Q9WKA7 human immun
59	41.7	113	14	Q9WKA6	Q9WKA6 human immun
60	41.7	113	14	Q9WKA5	Q9WKA5 human immun
61	41.7	113	14	Q9WKA4	Q9WKA4 human immun
62	41.7	113	14	Q9WKA3	Q9WKA3 human immun
63	41.7	113	14	Q9WKA2	Q9WKA2 human immun
64	41.7	113	14	Q9WKA1	Q9WKA1 human immun
65	41.7	113	14	Q9WKA8	Q9WKA8 human immun
66	41.7	113	14	Q9WKA7	Q9WKA7 human immun
67	41.7	113	14	Q9WKA6	Q9WKA6 human immun
68	41.7	113	14	Q9WKA5	Q9WKA5 human immun
69	41.7	113	14	Q9WKA4	Q9WKA4 human immun
70	41.7	113	14	Q9WKA3	Q9WKA3 human immun
71	41.7	113	14	Q9WKA2	Q9WKA2 human immun
72	41.7	113	14	Q9WKA1	Q9WKA1 human immun
73	41.7	113	14	Q9WKA8	Q9WKA8 human immun
74	41.7	113	14	Q9WKA7	Q9WKA7 human immun
75	41.7	113	14	Q9WKA6	Q9WKA6 human immun
76	41.7	113	14	Q9WKA5	Q9WKA5 human immun
77	41.7	113	14	Q9WKA4	Q9WKA4 human immun
78	41.7	113	14	Q9WKA3	Q9WKA3 human immun
79	41.7	113	14	Q9WKA2	Q9WKA2 human immun
80	41.7	113	14	Q9WKA1	Q9WKA1 human immun
81	41.7	113	14	Q9WKA8	Q9WKA8 human immun
82	41.7	113	14	Q9WKA7	Q9WKA7 human immun
83	41.7	113	14	Q9WKA6	Q9WKA6 human immun
84	41.7	113	14	Q9WKA5	Q9WKA5 human immun
85	41.7	113	14	Q9WKA4	Q9WKA4 human immun
86	41.7	113	14	Q9WKA3	Q9WKA3 human immun
87	41.7	113	14	Q9WKA2	Q9WKA2 human immun
88	41.7	113	14	Q9WKA1	Q9WKA1 human immun
89	41.7	113	14	Q9WKA8	Q9WKA8 human immun
90	41.7	113	14	Q9WKA7	Q9WKA7 human immun
91	41.7	113	14	Q9WKA6	Q9WKA6 human immun
92	41.7	113	14	Q9WKA5	Q9WKA5 human immun



**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (uspto)**

---

DR InterPro: IPR001245; -  
 DR Pfam: PF00017; SH2: 1.  
 DR Pfam: PF00069; PKinase: 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00001; SH2; 1.  
 DR Transferase: Tyrosine-protein kinase; ATP-binding; SH2 domain;  
 KM Phosphorylation.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 70 SH2.  
 FT DOMAIN 95 348 PROTEIN KINASE.  
 FT NP\_BIND 101 109 ATP (BY SIMILARITY).  
 FT BINDING 123 123 ATP (BY SIMILARITY).  
 FT ACT\_SITE 214 214 BY SIMILARITY.  
 SQ SEQUENCE 362 AA: 40937 MW: D8C549E904EB6D0 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 362;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPYA 5  
 |||||  
 DB 117 STPYA 121

RESULT 50  
 RRPO\_RECVD STANDARD; PRT; 366 AA.  
 ID RRPO\_RECVD  
 AC P03526;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (SIGMA NS PROTEIN) (CORE  
 DE PROTEIN) (RNA-BINDING PROTEIN).  
 GN S3.  
 OS Reovirus (type 3 / strain Deering).  
 OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.  
 OX NCBI\_TaxID=10886;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84015379; PubMed=6312421;  
 RA Richardson M.A., Furuichi Y.;  
 RT "Nucleotide sequence of reovirus genome segment S3, encoding non-  
 RT structural protein sigma NS.";  
 RL Nucleic Acids Res. 11:6399-6408(1983).  
 RN [2]  
 RP SEQUENCE OF 1-17 FROM N.A.  
 RA MEDLINE=83017876; PubMed=6927854;  
 RA Antczak J.B., Chmelo R.A., Pickup D.J., Joklik W.K.;  
 RT "Sequence at both termini of the 10 genes of reovirus serotype 3  
 RT (strain Deering)."  
 RL Virology 121:307-319(1982).  
 CC -1- FUNCTION: THIS PROTEIN HAS A POLY(C)-DEPENDENT POLY(G) POLYMERASE  
 CC ACTIVITY, AND BINDS SINGLE-STRANDED, BUT NOT DOUBLE-STRANDED, RNA.  
 CC -1- SIMILARITY: 86.3% SIMILARITY TO REOVIRUS SEROTYPE 2 SIGMA NS  
 CC PROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X01627; CAA25768.1; -  
 CC EMBL: J02330; AAA47280.1; -  
 CC PIR: A04126; MNXRSD.  
 DR InterPro: IPR002507; -  
 DR Pfam: PF01518; PolyG\_POL; 1.  
 KW Core protein; RNA-binding; Transferase; RNA-directed RNA polymerase.

SQ SEQUENCE 366 AA: 41056 MW: 201A6223C4290D7D CRC64;

Query Match 41.7%; Score 5; DB 1; Length 366;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
 |||||  
 DB 189 PVAPT 193

Search completed: August 15, 2001, 12:43:05  
 Job time: 495 sec

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X13464; CAA31824.1; -  
DR EMBL; J01804; AAA8616.1; -  
DR PIR; J0158; XNEBHC.  
DR StyGene; SG10159; hIsc.  
DR InterPro; IPR001917; -  
DR Pfam; PF00222; aminotran\_2.1.  
DR PROSITE; PS00599; AA\_TRANSFER\_CLASS\_2; 1.  
KW Histidine biosynthesis; Transferase; Aminotransferase;  
KW Pyridoxal phosphate  
FT BINDING 217 217 PYRIDOXAL PHOSPHATE (PROBABLE).  
FT CONFLICT 148 164 DGTAVFVCSPPNPTGQ -> TAQKWCSEVAPILPDN  
FT CONFLICT 226 226 C -> R (IN REF. 1).  
FT CONFLICT 260 260 L -> S (IN REF. 1).  
FT CONFLICT 264 265 GI -> ES (IN REF. 1).  
FT CONFLICT 349 349 R -> P (IN REF. 1).  
SQ SEQUENCE 359 AA; 39715 MW; 6241506A59CBB2A7 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 359;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STPVA 5  
|||||  
Db 249 STPVA 253

RESULT 48  
WNT2\_CAEEL  
ID WNT2\_CAEEL STANDARD: PRT; 360 AA.  
AC P34889;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE WNT-2 PROTEIN PRECURSOR.  
GN WNT-2 OR CMN-2 OR W01B6.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=93288400; PubMed=8510930;  
RA Shuckleford G.M., Shiva Kumar S., Shue L., Mason J., Kenyon C.,  
RA Varma H.E.;  
RT "Two wnt genes in Caenorhabditis elegans."  
RL Oncogene 8:1857-1864 (1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING  
CC MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF  
CC TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.  
CC -1- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE  
CC EXTRACELLULAR MATRIX.  
CC -1- DEVELOPMENTAL STAGE: DETECTED IN ALL LARVAL FORMS AND ADULTS,  
CC BUT IS MOST ABUNDANT IN THE EMBRYONIC STAGE.  
CC -1- SIMILARITY: BELONGS TO THE WNT FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X72943; CAA51448.1; -  
DR EMBL; Z68301; CAA92624.1; -  
DR PIR; S32695; S32695.  
DR MorphPep; M01B6.1; CE03753.  
DR InterPro; IPR000970; -  
DR Pfam; PF00110; wnt; 1.  
DR PROSITE; PS00246; WNT1; 1.  
KW Developmental protein; Glycoprotein; Signal.  
FT SIGNAL 1 37  
FT CHAIN 38 360 WNT-2 PROTEIN.  
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 73 73 R -> A (IN REF. 2).  
FT CONFLICT 186 186 D -> H (IN REF. 2).  
FT CONFLICT 221 221 R -> A (IN REF. 2).  
FT CONFLICT 229 230 TE -> MQ (IN REF. 2).  
FT CONFLICT 333 333 A -> S (IN REF. 2).  
SQ SEQUENCE 360 AA; 40551 MW; EA6EAF82A710F46 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 360;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TQEVK 11  
|||||  
Db 328 TQEVK 332

RESULT 49  
SRK2\_SPOLA  
ID SRK2\_SPOLA STANDARD: PRT; 362 AA.  
AC P42688;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE TYROSINE-PROTEIN KINASE SRK2 (EC 2.7.1.112) (FRAGMENT).  
GN SRK2.  
OS Spongilla lacustris (Freshwater sponge).  
OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;  
OC Haplosclerida; Spongillidae; Spongilla.  
OX NCBI\_TaxID=6055;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC MEDLINE=92334872; PubMed=1378585;  
RA Otfillie S., Raulf F., Barnekow A., Hanbig G., Scharf M.,  
RT "Multiple src-related kinase genes, srk1-4, in the fresh water sponge  
RT Spongilla lacustris."  
RL Oncogene 7:1625-1630 (1992).  
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +  
CC PROTEIN TYROSINE PHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).  
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
CC DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X61602; CAA43799.1; -  
DR HSSP; P00523; 2PKK.  
DR InterPro; IPR000719; -  
DR InterPro; IPR000980; -

OY 3 PVAPT 7  
 DB 284 PVAPT 288

RESULT 46  
 HIS8\_ECOLI  
 ID HIS8\_ECOLI STANDARD; PRT; 356 AA.  
 AC P06986;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (EC 2.6.1.9) (IMIDAZOLE ACETOL-  
 DE PHOSPHATE TRANSAMINASE).  
 OS HIS8.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OC NCBI\_Taxid=562;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RC MEDLINE=89094829; PubMed=3062174;  
 RA Carlomagno M.S., Chiaricelli L., Alfano P., Nappo A.G., Bruni C.B.;  
 RT "Structure and function of the Salmonella typhimurium and Escherichia  
 coli K-12 histidine operons.";  
 RL J. Mol. Biol. 203:585-606(1988).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RC MEDLINE=86059230; PubMed=2999081;  
 RA Gisolia V., Carlomagno M.S., Nappo A.G., Bruni C.B.;  
 RT "Cloning, structure, and expression of the Escherichia coli K-12 hisC  
 gene.";  
 RL J. Bacteriol. 164:1317-1323(1985).  
 RN [3]  
 RN REVISIONS, SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RC MEDLINE=94260549; PubMed=8201624;  
 RA Jovanovic G., Kostic T., Jankovic M., Savic D.J.;  
 RT "Nucleotide sequence of the Escherichia coli K12 histidine operon  
 revisited.";  
 RL J. Mol. Biol. 239:433-435(1994).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RC MEDLINE=97426517; PubMed=9278503;  
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RC MEDLINE=97251358; PubMed=9097040;  
 RA Itoh T., Alta H., Baba T., Fujita K., Hayashi K., Inada T.,  
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,  
 RA Saito N., Sempel G., Seki Y., Sivasubraman S., Tagami H.,  
 RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horinouchi T.,  
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome  
 corresponding to the 40,150,0 min region on the linkage map.";  
 RL DNA Res. 3:379-392(1996).  
 CC -1- CATALYTIC ACTIVITY: L-HISTIDINOL-PHOSPHATE + 2-OXOGlutARATE -  
 CC 3-(IMIDAZOL-4-YL)-2-OXOPROPYL PHOSPHATE + GLUTAMATE.  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -1- PATHWAY: EIGHTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT

CC AMINOTRANSFERASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X03416; CAA27150.1; -;  
 DR EMBL: X13462; CAA31813.1; -;  
 DR EMBL: U02071; AAA19743.1; -;  
 DR EMBL: AE000293; AAC75082.1; -;  
 DR EMBL: D90840; CAB21779.1; -;  
 DR PIR: A30270; XNECHC.  
 DR EcoGene: EG10446; HIS8.  
 DR InterPro: IPR001917; -;  
 DR Pfam: PF00222; aminotran\_2; 1.  
 DR PROSITE: PS00599; AA\_TRANSF\_CLASS\_2; 1.  
 KW Histidine biosynthesis; Transferase; Aminotransferase;  
 KW Pyridoxal phosphate.  
 FT BINDING 214 214 PYRIDOXAL PHOSPHATE (PROBABLE).  
 FT CONFLICT 130 130 L -> P (IN REF. 1 AND 2).  
 FT CONFLICT 149 149 V -> A (IN REF. 1 AND 2).  
 SQ SEQUENCE 356 AA; 39360 MW; 056CB3C8F894083F CRC64;

Query Match 41.7%; Score 5; DB 1; Length 356;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVA 5  
 DB 246 STPVA 250

RESULT 47  
 HIS8\_SALTY  
 ID HIS8\_SALTY STANDARD; PRT; 359 AA.  
 AC P10369;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-FEB-1995 (Rel. 31, Last annotation update)  
 DE HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (EC 2.6.1.9) (IMIDAZOLE ACETOL-  
 DE PHOSPHATE TRANSAMINASE).  
 OS HIS8.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OC NCBI\_Taxid=602;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-LT2;  
 RC MEDLINE=89094829; PubMed=3062174;  
 RA Carlomagno M.S., Chiaricelli L., Alfano P., Nappo A.G., Bruni C.B.;  
 RT "Structure and function of the Salmonella typhimurium and Escherichia  
 coli K-12 histidine operons.";  
 RL J. Mol. Biol. 203:585-606(1988).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-LT2;  
 RC MEDLINE=97251358; PubMed=9097040;  
 RA Itoh T., Alta H., Baba T., Fujita K., Hayashi K., Inada T.,  
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,  
 RA Saito N., Sempel G., Seki Y., Sivasubraman S., Tagami H.,  
 RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horinouchi T.,  
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome  
 corresponding to the 40,150,0 min region on the linkage map.";  
 RL DNA Res. 3:379-392(1996).  
 CC -1- CATALYTIC ACTIVITY: L-HISTIDINOL-PHOSPHATE + 2-OXOGlutARATE -  
 CC 3-(IMIDAZOL-4-YL)-2-OXOPROPYL PHOSPHATE + GLUTAMATE.  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -1- PATHWAY: EIGHTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
 CC AMINOTRANSFERASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

RA Deschamps S., Viel A., Denis H., le Maire M.;  
RT "Purification of two thermostable components of messenger  
CC ribonucleoprotein particles (mRNPs) from *Xenopus laevis* oocytes,  
CC belonging to a novel class of RNA-binding proteins.";  
RL FEBS Lett. 282:110-114(1991).  
CC  
CC -1- FUNCTION: BINDS TO CCAT-CONTAINING Y BOX OF THE HSP70 GENES.  
CC SEEMS TO BE A NEGATIVE REGULATORY FACTOR. ALSO BINDS TO MRNA.  
CC  
CC -1- SUBUNIT: POSSIBLY FORMS A HETERODIMER WITH P54 IN THE 6S AND 15S  
CC MRNA-BINDING PARTICLES.  
CC  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, EITHER FREE OR ASSOCIATED WITH  
CC RIBONUCLEOPROTEIN PARTICLES.  
CC  
CC -1- TISSUE SPECIFICITY: TESTIS AND IMMATURE OOCYTES.  
CC  
CC -1- PM: PHOSPHORYLATION ACTIVATES IN VITRO RNA-BINDING.  
CC  
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC  
CC EMBL; M59454; AAA49716.1; -  
CC PIR; B38274; B38274.  
CC HSSP; P32081; INMG.  
CC TRASNFPAC; T00294; -  
CC InterPro; IPR002059; -  
CC Pfam; PF00313; CSD; 1.  
CC PRINTS; PR00050; COLDSHOCK.  
CC DR PROSITE; PS00352; COLD\_SHOCK; 1.  
CC KW Transcription regulation; DNA-binding; Nuclear protein;  
KW RNA-binding; Phosphorylation.  
FT DOMAIN 44 108 CSD.  
FT CONFLICT 254 254 A -> T (IN REF. 1).  
FT SEQUENCE 336 AA: 37202 MW: 4AD5838769C6B84D CRC64;  
SQ  
  
Query Match 41.7%; Score 5; DB 1; Length 336;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC  
CC -1- PATHWAY: LAST (FIFTH) STEP IN BIOSYNTHESIS OF TRYPTOPHAN.  
CC  
CC -1- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS.  
CC  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST (PROBABLE).  
CC  
CC -1- SIMILARITY: BELONGS TO THE TRPA FAMILY.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC  
CC EMBL; X76713; CA54131.1; -  
CC HSSP; P00929; TMSV.  
CC MaizeDB; 102199; -  
CC InterPro; IPR002028; -  
CC Pfam; PF00290; trp-synta; 1.  
CC DR PROSITE; PS00167; TRP SYNTHASE ALPHA; 1.  
CC KW Tryptophan biosynthesis; Lyase; Chloroplast; Transit peptide.  
FT TRANSIT 1 346 CHLOROPLAST (POTENTIAL).  
FT CHAIN 1 346 TRYPTOPHAN SYNTHASE ALPHA CHAIN.  
FT SEQUENCE 346 AA: 36967 MW: 05F8FC2635071F3E CRC64;  
SQ  
  
Query Match 41.7%; Score 5; DB 1; Length 346;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 GEVKK 12  
Db 278 GEVKK 262

RESULT. 45  
YMA3\_MYCBO STANDARD; PRT; 354 AA.  
ID YMA3\_MYCBO  
AC 002279;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE HYPOTHELICAL 38.1 KDA PROTEIN IN MAS 3/REGION.  
OS Mycobacterium bovis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1765;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCG;  
RX MEDLINE=92406887; PubMed=1527058;  
RA Mathur M., Kolattukudy P.E.;  
RT "Molecular cloning and sequencing of the gene for mycroceroic acid  
RT synthase, a novel fatty acid elongating multifunctional enzyme, from  
RT Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guerin.";  
RL J. Biol. Chem. 267:19388-19395(1992).  
CC  
CC -1- SIMILARITY: SOME TO M.TUBERCULOSIS RV2285.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC  
CC EMBL; M95808; AAA25370.1; -  
CC KW Hypothetical protein.  
FT SEQUENCE 354 AA: 37956 MW: 2C77C1259BD3686A CRC64;  
SQ

Query Match 41.7%; Score 5; DB 1; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

CC MODABCD FOR MOLYBDENUM.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE MODD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X80268; CAA56555.1; -
DR EMBL: X99258; CAA67645.1; -
DR EMBL: Z83859; CAB06127.1; -
DR Tuberculist; RY1860; -
KM Molybdenum; Transport; Glycoprotein; Antigen; Signal.
FT SIGNAL 1 39
FT CHAIN 40 325 MOLYBDATE UPTAKE SECRETED PROTEIN.
FT CARBOHYD 49 49 O-LINKED (MAN. . .)
FT CARBOHYD 57 57 O-LINKED (MAN. . .)
FT CARBOHYD 66 66 O-LINKED (MAN. . .)
FT CARBOHYD 316 316 O-LINKED (MAN. . .)
SQ SEQUENCE 325 AA; 32720 MM; 59ESD0455A97BED CRC64;

Query Match 41.7%; Score 5; DB 1; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6
Db 66 TPVAP 70

RESULT 42
ENT2_MOUSE STANDARD; PRT; 327 AA.
ID ENT2_MOUSE 061672;
AC 061672;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE EQUILIBRATIVE NUCLEOSIDE TRANSPORTER 2 (EQUILIBRATIVE
DE NITROBENZYLAMERAPTOPURINE RIBOSIDE-INSENSITIVE NUCLEOSIDE TRANSPORTER)
DE (EQUILIBRATIVE NBMPR-INSENSITIVE NUCLEOSIDE TRANSPORTER) (NUCLEOSIDE
DE TRANSPORTER, ET-TYPE) (36 KDA NUCLEOLAR PROTEIN HNP36) (HYDROPHOBIC
DE NUCLEOLAR PROTEIN, 36 KDA) (DELAYED-EARLY RESPONSE PROTEIN 12).
GN SLC29A2 OR ENT2 OR HNP36 OR DER12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathii; Muridae; Murinae; Mus.
OX NCBI_Taxid=10950;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN-BALB/C; TISSUE-Fibroblast;
RX MEDLINE-95367016; PubMed-7639753;
RA Williams J.B., Lananah A.A.;
RT "A mammalian delayed-early response gene encodes HNP36, a novel,
RT conserved nucleolar protein."
RL Biochem. Biophys. Res. Commun. 213:325-333(1995).
CC -1- FUNCTION: MEDIATES EQUILIBRATIVE TRANSPORT OF PURINE AND
CC PYRIMIDINE NUCLEOSIDES, AND THE PURINE BASE HYPOXANTHINE.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR. INTEGRAL MEMBRANE
CC PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM AND A SHORT
CC FORM/HNP36 (SHOWN HERE); SEEMS TO BE PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- INDICATION: BY PLATELET DERIVED GROWTH FACTOR (PDGF) AND FIBROBLAST
CC GROWTH FACTOR (FGF).
CC -1- SIMILARITY: BELONGS TO THE SLC29A FAMILY OF TRANSPORTERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X86682; CAA60381.1; -
DR MGD: MG1:1345278; SLC29a2.
DR InterPro: IPR002259; -
DR Pfam: PF01733; Nucleoside_tran; 1.
DR PRINTS: PR01130; DERENTRNSPT.
KM Nuclear protein; Transmembrane; Transport; Alternative splicing.
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 32 52 POTENTIAL.
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 159 179 POTENTIAL.
FT TRANSMEM 194 214 POTENTIAL.
FT TRANSMEM 231 251 POTENTIAL.
FT TRANSMEM 267 287 POTENTIAL.
FT TRANSMEM 303 323 POTENTIAL.
SQ SEQUENCE 327 AA; 36113 MM; 5D2D3FF4BBD592B6 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 327;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APTOE 9
Db 100 APTOE 104

RESULT 43
YB56_XENLA STANDARD; PRT; 336 AA.
ID YB56_XENLA P21574;
AC P21574;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CYTOPLASMIC RNA-BINDING PROTEIN P56 (Y BOX BINDING PROTEIN-2) (Y-BOX
DE TRANSCRIPTION FACTOR) (MRNP4).
GN FRG2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OX Xenopodinae; Xenopus.
OX NCBI_Taxid=8335;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE-91062413; PubMed-2247479;
RA Tafuri S.R., Wolfe A.P.;
RT "Xenopus Y-box transcription factors: molecular cloning, functional
RT analysis and developmental regulation."
RL Proc. Natl. Acad. Sci. U.S.A. 87:9028-9032(1990).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-92107999; PubMed-1729676;
RA Murray M.T., Schiller D.L., Franke W.W.;
RT "Sequence analysis of cytoplasmic mRNA-binding proteins of Xenopus
RT oocytes identifies a family of RNA-binding proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 89:11-15(1992).
RN [3]
RP PARTIAL SEQUENCE.
RC TISSUE-Ovary;
RX MEDLINE-92332467; PubMed-1629179;
RA Deschamps S., Viel A., Garrigos M., Denis H., le Maire M.;
RT "MRNP4, a major mRNA-binding protein from Xenopus oocytes is
RT identical to transcription factor FRG Y2."
RL J. Biol. Chem. 267:13799-13802(1992).
RN [4]
RP PARTIAL SEQUENCE.
RC TISSUE-Ovary;
RX MEDLINE-91224309; PubMed-1902800;

```

DR MGD; MGI:894407; Tpar1.  
 DR InterPro; IPR001727; .  
 DR Pfam; PF01169; UPE0016; 1.  
 DR PROSITE; PS01214; UPE0016; 1.  
 KW Transmembrane.  
 FT TRANSMEM 16 36 POTENTIAL.  
 FT TRANSMEM 90 110 POTENTIAL.  
 FT TRANSMEM 127 147 POTENTIAL.  
 FT TRANSMEM 152 172 POTENTIAL.  
 FT TRANSMEM 228 248 POTENTIAL.  
 FT TRANSMEM 267 287 POTENTIAL.  
 FT TRANSMEM 299 319 POTENTIAL.  
 SO SEQUENCE 323 AA; 34693 MM; 4CID1201ADEFL333 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
 |||||  
 DB 73 TPVAP 77

RESULT 40  
 MODD\_MYCBO STANDARD; PRT: 325 AA.  
 AC 030620; 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE MOLYBDATE UPTAKE SECRETED PROTEIN PRECURSOR (ALA-PRO RICH 45/47 KDA  
 DE SECRETED PROTEIN) (IMMUNOGENIC PROTEIN MPT3?) (ANTIGEN MPT-32) (45-KDA  
 DE GLYCOPROTEIN) (FIBRONECTIN ATTACHMENT PROTEIN) (FAP-B).  
 GN MODD OR APA.  
 OS Mycobacterium bovis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1765;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=BCG;  
 RA Zhao W., Schorey J.S., Bong-Mastek M., Brown E.J., Ratliff T.L.;  
 RT "Identification, sequence and characterization of the M. bovis BCG  
 RT fibronectin attachment protein.";  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM  
 CC MODABD FOR MOLYBDENUM (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO THE MODD FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
 CC -----  
 DR EMBL; AF013569; AAB71842.1; .  
 KW Molybdenum; Transport; Glycoprotein; Antigen; Signal.  
 FT SIGNAL 1 39 BY SIMILARITY.  
 FT CHAIN 40 325 MOLYBDATE UPTAKE SECRETED PROTEIN.  
 FT CARBOHYD 49 49 O-LINKED (MAN. . .) (BY SIMILARITY).  
 FT CARBOHYD 57 57 O-LINKED (MAN. . .) (BY SIMILARITY).  
 FT CARBOHYD 66 66 O-LINKED (MAN. . .) (BY SIMILARITY).  
 FT CARBOHYD 316 316 O-LINKED (MAN. . .) (BY SIMILARITY).  
 SO SEQUENCE 325 AA; 32686 MM; D3419CA5547D91E9 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 325;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
 |||||  
 DB 66 TPVAP 70

RESULT 41  
 MODD\_MYCTU STANDARD; PRT: 325 AA.  
 AC 050906; 008062; 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE MOLYBDATE UPTAKE SECRETED PROTEIN PRECURSOR (ALA-PRO RICH 45/47 KDA  
 DE SECRETED PROTEIN) (IMMUNOGENIC PROTEIN MPT3?) (ANTIGEN MPT-32) (45-KDA  
 DE GLYCOPROTEIN) (FIBRONECTIN ATTACHMENT PROTEIN) (FAP-B).  
 GN MODD OR APA OR RV1860 OR MTCY359.13.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=H37RV;  
 RC MEDLINE=96009758; PubMed=7558311;  
 RA Laqueyrie A., Miltzer P., Romain F., Eiglmeyer K., Cole S.,  
 RA Marchal G.;  
 RT "Cloning, sequencing, and expression of the apa gene coding for the  
 RT Mycobacterium tuberculosis 45/47-kilodalton secreted antigen  
 RT complex.";  
 RL Infect. Immun. 63:4003-4010(1995).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=H37RV;  
 RA Laqueyrie A.;  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=H37RV;  
 RC MEDLINE=96295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Pettwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,  
 RA Taylor K., Whitehead S., Barrett B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 RN [4]  
 RP PARTIAL SEQUENCE, AND CHARACTERIZATION.  
 RX MEDLINE=9109989; PubMed=1898899;  
 RA Nagai S., Wiker H.G., Harboe M., Kinomoto M.;  
 RT "Isolation and partial characterization of major protein antigens in  
 RT the culture fluid of Mycobacterium tuberculosis.";  
 RL Infect. Immun. 59:372-382(1991).  
 RN [5]  
 RP PARTIAL SEQUENCE, AND GLYCOSYLATION.  
 RX MEDLINE=9547792; PubMed=7622204;  
 RA Dobos K.M., Swiderex K., Khoo K.-H., Brennan P.J., Bellisle J.T.;  
 RT "Evidence for glycosylation sites on the 45-kilodalton glycoprotein of  
 RT Mycobacterium tuberculosis.";  
 RL Infect. Immun. 63:2846-2853(1995).  
 RN [6]  
 RP PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.  
 RX MEDLINE=96196153; PubMed=8626314;  
 RA Dobos K.M., Khoo K.-H., Swiderex K.M., Brennan P.J., Bellisle J.T.;  
 RT "Definition of the full extent of glycosylation of the 45-kilodalton  
 RT glycoprotein of Mycobacterium tuberculosis.";  
 RL J. Bacteriol. 178:2498-2506(1996).  
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM



RX MEDLINE-93250487; PubMed-7763615;  
 RA Dupille E., Rombaldi C., Lelievre J.M., Cleyet-Marel J.C., Pech J.C.,  
 RA Latche A.;  
 RT "Purification, properties and partial amino-acid sequence of 1-  
 RT aminocyclopropane-1-carboxylic acid oxidase from apple fruits.";  
 RL Planta 190;65-70(1993).  
 CC -1- CATALYTIC ACTIVITY: 1-AMINOCYCLOPROPANE-1-CARBOXYLATE + O2 =  
 CC ETHYLENE + HCN + CO(2) + 2 H(2)O.  
 CC -1- COFACTOR: IRON AND ASCORBATE.  
 CC -1- PATHWAY: LAST STEP IN THE BIOSYNTHESIS OF ETHYLENE.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING FRUIT RIPENING.  
 CC -1- INDUCTION: BY ETHYLENE AND BY WOUNDING.  
 CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF  
 CC OXIDOREDUCTASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X61390; CAA33662.1; -;  
 CC EMBL: M61794; AAA33412.1; -;  
 CC EMBL: Y14005; CAA74328.1; -;  
 CC PIR: S22513; S22513.  
 CC DR InterPro: IPR002419; -;  
 CC DR Pfam: PF00671; Fe\_Asc\_Oxidored; 1.  
 CC DR Fruit Ripening; Ethylene biosynthesis; Oxidoreductase; Iron;  
 CC KW Vitamin C; Multigene family.  
 CC SQ SEQUENCE 314 AA; 35410 MW; BEADA64C7AD10E1E CRC64;  
  
 Query Match 41.7%; Score 5; DB 1; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 OY 1 STPYA 5  
 Db 308 STPYA 312  
  
 RESULT 38  
 ID SPLR.NPVO P STANDARD; PRT; 321 AA.  
 AC 065328; 012842; 012553;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE SPHEROIDIN-LIKE PROTEIN PRECURSOR (SPINDOLIN-LIKE PROTEIN) (37 KDA  
 DE GLYCOPROTEIN).  
 GN SLR OR GP37.  
 OS Orygia pseudotsugata multicausid polyhedrosis virus (OpNPV).  
 OS Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 OC Nucleopolyhedrovirus.  
 CC NCBI\_TaxID=10450;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE-93100831; PubMed-8380088;  
 RA Gross C.H., Wolgamot G.M., Russell R.L., Pearson M.N., Rohrmann G.F.;  
 RT "A 37-kilodalton glycoprotein from a baculovirus of Orygia  
 RT pseudotsugata is localized to cytoplasmic inclusion bodies.";  
 RL J. Virol. 67:469-475(1993).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE-97271300; PubMed-9126251;  
 RA Atkins C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,  
 RA Rohrmann G.F.;  
 RT "The sequence of the Orygia pseudotsugata multicausid nuclear  
 RT polyhedrosis virus genome.";  
 RL Virology 229:381-399(1997).

CC -1- FUNCTION: COMPONENT OF THE VIRUS OCCLUSION BODIES, WHICH ARE LARGE  
 CC PROTEINACEOUS STRUCTURES (POLYHEDRA), THAT PROTECT THE VIRUS FROM  
 CC THE OUTSIDE ENVIRONMENT FOR EXTENDED PERIODS UNTIL THEY ARE  
 CC INGESTED BY INSECT LARVAE (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE PERIPHERY OF OCCLUSION  
 CC BODIES AND WITH THE INTERNAL MEMBRANES OF INFECTED CELLS.  
 CC -1- DEVELOPMENTAL STAGE: VERY LATE PHASE OF INFECTION.  
 CC -1- SIMILARITY: TO CBEPV AND HAPV SPINDOLINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D13306; BAA02566.1; -;  
 CC EMBL: U75930; AAC59068.1; -;  
 CC KW Viral occlusion body; Signal; Glycoprotein; Late protein.  
 CC FT SIGNAL 1 18 POTENTIAL.  
 CC FT CHAIN 19 321 SPHEROIDIN-LIKE PROTEIN.  
 CC FT CARBOHYD 192 192 N-LINKED (GLCNAC...) (POTENTIAL).  
 CC SQ SEQUENCE 321 AA; 36136 MW; B935809F06B56CA9 CRC64;  
  
 Query Match 41.7%; Score 5; DB 1; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 OY 1 STPYA 5  
 Db 23 STPYA 27  
  
 RESULT 39  
 ID PF27.MOUSE STANDARD; PRT; 323 AA.  
 AC P52875;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE TRANSMEMBRANE PROTEIN PF27 (TPA REGULATED LOCUS PROTEIN).  
 GN TPRL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC TISSUE-Teratocarcinoma;  
 RX MEDLINE-89076283; PubMed-3202867;  
 RA Akagi J., Nomiyama H., Setoyama C., Shimada K., Akagi M.;  
 RT "Messenger RNA expressed in mouse teratocarcinoma stem cells and  
 RT down-regulated by a tumor-promoting phorbol ester codes for a novel  
 RT transmembrane protein.";  
 RL Biochem. Biophys. Res. Commun. 157:548-557(1988).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN UNDIFFERENTIATED MOUSE F9  
 CC TERATOCARCINOMA STEM CELLS BUT DISAPPEARING RAPIDLY AFTER  
 CC TREATMENT WITH A TUMOR-PROMOTING PHORBOL ESTER.  
 CC -1- SIMILARITY: BELONGS TO THE UPF0016 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M23568; AAA40456.1; -;

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M63743; AAA33997.1; -  
 DR EMBL: M95400; AAA34019.1; -  
 DR EMBL: M00353; AAA33995.1; -  
 DR EMBL: D86929; BAA13184.1; -  
 DR EMBL: D86930; BAA13185.1; -  
 DR EMBL: AB002810; BAA19673.1; -  
 DR PIR: A25776; A25776.  
 DR Mendel: 13686; GLYma:1087.1.  
 DR Mendel: 16777; GLYma:1087.mm16777.  
 DR InterPro: IPR002042; -  
 DR Pfam: PF01014; URICASE; 1.  
 DR PRINTS: PR00093; URICASE.  
 DR PROSITE: PS00366; URICASE; 1.  
 DR PROSITE: PS00342; MICROBODIES\_CTER; 1.  
 DR Modulation: Oxidoreductase; Peroxisome; Purine metabolism.  
 FT SITE 307 309 MICROBODY TARGETING SIGNAL (POTENTIAL).  
 FT VARIANT 3 3 Q -> K (IN CV. AKISENGOKU).  
 FT VARIANT 235 235 S -> T.  
 FT CONFLICT 60 60 D -> E (IN REF. 2).  
 FT CONFLICT 195 195 Y -> C (IN REF. 4).  
 SQ SEQUENCE 309 AA: 35138 MW: 6FBC2A758FBA982 CRC64:

Query Match 41.7%; Score 5; DB 1; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 1,2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEVRK 12  
 Db 215 OEVRK 219

RESULT 36  
 ID URID\_SOYBN STANDARD: PRT: 309 AA.  
 AC 004104;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DE URICASE II ISOZYME 2 (EC 1.7.3.3) (URATE OXIDASE) (NODULIN 35) (N-35)  
 DE (NON-SYMBIOTIC URICASE).  
 OS Glycine max (Soybean).  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 CC Fabales; Fabaceae; Papilionoideae; Glycine.  
 CC NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 245-262.  
 RC STRAIN-CV, DARE:  
 RA Nguyen T., Zelechowska M., Foster V., Bergmann H., Verma D.P.S.,  
 RT "Primary structure of the soybean nodulin-35 gene encoding uricase II  
 RT localized in the peroxisome of uninfected cells of nodules";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:5040-5044(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV, AKISENGOKU; TISSUE=Root nodules;  
 RA Takane K., Tajima S., Kouchi H.;  
 RT "Two distinct uricase II (nodulin 35) genes are differentially  
 RT expressed in soybean plants";  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O = 5-HYDROXYISOURATE +  
 CC H(2)O(2) (5-HYDROXYISOURATE DECOMPOSE TO FORM ALLANTOIN).  
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: PEROXISOMAL.  
 CC -1- SIMILARITY: BELONGS TO THE URICASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M10594; AAA33994.1; -  
 DR EMBL: AB002807; BAA19670.1; -  
 DR EMBL: AB002809; BAA19672.1; -  
 DR Mendel: 16776; GLYma:1087.mm16776.  
 DR InterPro: IPR002042; -  
 DR Pfam: PF01014; URICASE; 1.  
 DR PRINTS: PR00093; URICASE.  
 DR PROSITE: PS00366; URICASE; 1.  
 DR PROSITE: PS00342; MICROBODIES\_CTER; 1.  
 DR Oxidoreductase; Peroxisome; Purine metabolism.  
 FT CONFLICT 208 208 L -> F (IN REF. 1).  
 SQ SEQUENCE 309 AA: 35137 MW: 81B8F87ACAA00A9 CRC64:

Query Match 41.7%; Score 5; DB 1; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 1,2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEVRK 12  
 Db 215 OEVRK 219

RESULT 37  
 ID ACCO\_MALDO STANDARD: PRT: 314 AA.  
 AC 000985; 024063;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE (EC 1.-.-.-) (ACC OXIDASE)  
 DE (ETHYLENE-FORMING ENZYME) (EFE) (PROTEIN AP4 OR PAE12).  
 OS Malus domestica (Apple) (Malus sylvestris).  
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 CC OC Rosales; Rosaceae; Malus.  
 CC NCBI\_TaxID=3750;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV, GOLDEN DELICIOUS; TISSUE=Fruit;  
 RX MEDLINE=92322950; PubMed=137761;  
 RA Ross G.S., Knighton M.L., Lay-Yee M.;  
 RT "An ethylene-related cDNA from ripening apples";  
 RL Plant Mol. Biol. 19:231-238(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Dong J.G., Olsen D.B., Silverstone A., Yang S.F.;  
 RT "Sequence of a cDNA coding for a 1-aminocyclopropane-1-carboxylate  
 RT oxidase homolog from apple fruit";  
 RL Plant Physiol. 98:1530-1531(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV, GOLDEN DELICIOUS;  
 RA Castiglione S., Malerba M., Pirolo B., Blanchetti R., Sala F.,  
 RL Ventura M., Pancaldi M., Sansavini S.;  
 RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP CHARACTERIZATION, AND SEQUENCE OF 115-134.  
 RX MEDLINE=93028548; PubMed=1409700;  
 RA Dong J.G., Fernandez-Maculet J.C., Yang S.F.;  
 RT "Purification and characterization of  
 RT 1-aminocyclopropane-1-carboxylate oxidase from apple fruit";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:9789-9793(1992).  
 RN [5]  
 RP CHARACTERIZATION, AND PARTIAL SEQUENCE.

CC -1- PATHWAY: FIRST STEP IN THE BIOLUMINESCENT FATTY ACID REDUCTASE  
 CC COMPLEX.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X65612; CAA46561.1; -  
 CC PIR: JN0518; JN0518.  
 CC HSSP: P05521; 1THT.  
 CC Luminescence: Transferase: Acyltransferase.  
 CC SSEQUENCE 305 AA; 34384 MW; 0BDAA76D08913164 CRC64;  
 SO

Query Match 41.7%; Score 5; DB 1; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVAK 12  
 Db 219 QEVAK 223

RESULT 34  
 POLG\_HCVH7 STANDARD: PRT; 309 AA.  
 AC P27955;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE GENOME POLYPROTEIN (CONTAINS: MATRIX PROTEIN (ENVELOPE PROTEIN M);  
 DE MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL PROTEIN NS1) (FRAGMENT).  
 OS Hepatitis C virus (isolate HCT27) (HCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11109;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91112009; PubMed=1846505;  
 RA Weiner A.J., Brauer M.J., Rosenblatt J., Richman K.H., Tung J.,  
 RA Crawford K., Bonino F., Saracco G., Choo Q.-L., Houghton M.,  
 RA Han J.H.;  
 RT "Variable and hypervariable domains are found in the regions of HCV  
 RT corresponding to the flavivirus envelope and NS1 proteins and the  
 RT pestivirus envelope glycoproteins";  
 RL Virology 180:842-848(1991).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X53133; CAA37293.1; -  
 CC InterPro: IPR002519; -  
 CC InterPro: IPR002531; -  
 CC InterPro: IPR002531; -  
 CC Pfam: PF01560; HCV\_NSI; 1.  
 CC Pfam: PF01562; HCV\_core; 1.  
 CC Pfam: PF01559; HCV\_env; 1.  
 CC Polyprotein: Glycoprotein; Coat protein; Envelope protein;  
 CC Transmembrane; Nonstructural protein.  
 CC Non\_TER 1  
 FT CHAIN <1 63 MATRIX PROTEIN (POTENTIAL).  
 FT CHAIN 64 255 MAJOR ENVELOPE PROTEIN E (POTENTIAL).  
 FT CHAIN 256 >309 NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC

FT CARBOHYD 81 81 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT NON\_TER 309 309  
 SO SEQUENCE 309 AA; 32922 MW; 6E858E9C3D0B9EA9 CRC64;  
 DB 113 PVAPT 117

Query Match 41.7%; Score 5; DB 1; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
 Db 113 PVAPT 117

RESULT 35  
 URIC\_SOYBN STANDARD: PRT; 309 AA.  
 ID URIC\_SOYBN  
 AC P04670; P34805; P93160; P93161; O04105;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE URICASE II ISOZYME 1 (EC 1.7.3.3) (URATE OXIDASE) (NODULIN 35) (N-35)  
 DE (NODULE SPECIFIC URICASE).  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;  
 OC Fabales; Fabaceae; Papilionoideae; Glycine.  
 OX NCBI\_TaxID=3847;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-CV. PRIZE;  
 RC Suzuki H., Verma D.P.S.;  
 RT "Soybean nodule-specific uricase (nodulin-35) is expressed and  
 RT assembled into a functional tetrameric holoenzyme in Escherichia  
 RT coli";  
 RL Plant Physiol. 95:384-389(1991).  
 CC -----  
 CC SEQUENCE FROM N.A. AND SEQUENCE OF 245-262.  
 CC STRAIN-CV. PRIZE;  
 RA Nguyen T., Zelechowska M., Foster V., Bergmann H., Verma D.P.S.;  
 RT "Primary structure of the soybean nodulin-35 gene encoding uricase II  
 RT localized in the peroxisome of uninfected cells of nodules";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:5040-5044(1985).  
 CC -----  
 CC SEQUENCE FROM N.A.  
 RA STRAIN-CV. AKISENGOKU. TISSUE-cotyledon, and Root nodules;  
 RA Takane K., Tanaka K., Tajima S., Okazaki K., Kouchi H.;  
 RT "Expression of uricase II gene (nodulin-35) in cotyledons of soybean  
 RT plants";  
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC SEQUENCE FROM N.A.  
 RA STRAIN-CV. AKISENGOKU. TISSUE-Root nodules;  
 RA Takane K., Tajima S., Kouchi H.;  
 RT "Two distinct uricase II (nodulin 35) genes are differentially  
 RT expressed in soybean plants";  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC FUNCTION: THIS IS A SUBUNIT OF THE NODULE SPECIFIC URICASE.  
 CC -1- CATALYTIC ACTIVITY: URATE + O(2) + H(2)O -> 5-HYDROXYISOURATE +  
 CC H(2)O(2) (5-HYDROXYISOURATE DECOMPOSE TO FORM ALLANTOIN).  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: PEROXISOMAL.  
 CC -1- INDUCTION: DURING NODULATION IN LEGUME ROOTS AFTER RHIZOBIUM  
 CC INFECTION.  
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.  
 CC -1- SIMILARITY: BELONGS TO THE URICASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 268 AA; 30386 MW; 2746ED6CC5D44AF2 CRC64;  
 Query Match 41.7%; Score 5; DB 1; Length 268;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 STPVA 5  
 11111  
 26 STPVA 30  
 RESULT 31  
 URED\_SYNY3  
 ID URED\_SYNY3 STANDARD: PRT: 270 AA.  
 AC P73047;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE UREASE ACCESSORY PROTEIN URED.  
 GN URED OR SLI1639.  
 OS Synchocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.  
 OX NCBI\_TaxID=1148;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneo T., Sato S., Kocani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirosewa M., Sugitani M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraaki A., Nakazaki N., Nario K., Okumura S.,  
 RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yasuda M.,  
 RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 Synchocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 CC -1- FUNCTION: PROBABLY FACILITATES NICKEL INCORPORATION.  
 CC -1- SIMILARITY: BELONGS TO THE URED FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D90903; BAA17068.1; ALT\_INIT.  
 DR InterPro: IPR002669; -  
 DR Pfam: PF01774; Ured; 1.  
 KW Nickel.  
 SQ SEQUENCE 270 AA; 30522 MW; FAA820E257E69286 CRC64;  
 Query Match 41.7%; Score 5; DB 1; Length 270;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 7 TOEVA 11  
 11111  
 254 TOEVA 258  
 RESULT 32  
 YBZ7\_YEAST  
 ID YBZ7\_YEAST STANDARD: PRT: 296 AA.  
 AC P38279;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE HYPOTHEETICAL 33.5 KDA PROTEIN IN MRPS9-YSM1 INTERGENIC REGION.  
 GN YBR147W OR YBR1124.

OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=5288C;  
 RA Enlian K.-D., Koetter P., Rose M., Becker J., Grey M., Li Z.,  
 RA Niegemann E., Schenk-Groeninger R., Servos J., Wehner E.,  
 RA Wolter R., Brendel M., Bauer J., Braun H., Derr K., Duesterhus S.,  
 RA Gruenbein R., Hedges D., Klesau P., Korol S., Krebs B., Proft M.,  
 RA Siegers K., Baur A., Boles E., Miosga T.,  
 RA Schaeft-Gerstenschlaeger I., Zimmermann F.K.;  
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -1- SIMILARITY: SOME, TO S.POMBE SPAC2E12.03C AND SPAC17C9.10.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z36016; CA85105.1; -  
 DR PIR: S46018; S46018.  
 DR SGD: S0000351; YBR147W.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 13 33 POTENTIAL.  
 FT TRANSMEM 45 65 POTENTIAL.  
 FT TRANSMEM 69 89 POTENTIAL.  
 FT TRANSMEM 164 184 POTENTIAL.  
 FT TRANSMEM 200 220 POTENTIAL.  
 FT TRANSMEM 239 259 POTENTIAL.  
 FT TRANSMEM 263 283 POTENTIAL.  
 SQ SEQUENCE 296 AA; 33496 MW; 50D65896A1BAFB43 CRC64;  
 Query Match 41.7%; Score 5; DB 1; Length 296;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 8 QEVK 12  
 11111  
 103 QEVK 107  
 RESULT 33  
 LXD2\_PHOE  
 ID LXD2\_PHOE STANDARD: PRT: 305 AA.  
 AC Q06878;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE ACYL TRANSFERASE (EC 2.3.1.-) (ACT) (MYRISTOYL-ACP-SPECIFIC  
 DE THIOESTERASE).  
 GN LUXD.  
 OS Photobacterium leiognathi.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;  
 OC Photobacterium.  
 OX NCBI\_TaxID=658;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=741;  
 RX MEDLINE=93231531; PubMed=8472957;  
 RA Chao Y.-F., Weng S.-F., Lin J.-W.;  
 RT "Sequence of the luxD gene encoding acyltransferase of the lux operon  
 RT from Photobacterium leiognathi.";  
 RL Gene 126:155-156(1993).  
 CC -1- FUNCTION: ACYL TRANSFERASE IS PART OF THE FATTY ACID REDUCTASE  
 CC SYSTEM REQUIRED FOR ALDEHYDE BIOSYNTHESIS. IT PRODUCES FATTY  
 CC ACIDS FOR THE LUMINESCENT REACTION.

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-L2/434/BU: PLASMID-PLGV440;
RX MEDLINE-89013895; PubMed-2845228;
RA Comanducci M., Ricci S., Ratti G.;
RT "The structure of a plasmid of Chlamydia trachomatis believed to be
RL required for growth within mammalian cells.";
RM Mol. Microbiol. 2:531-538(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-L1/440/LN: PLASMID-PLGV440;
RX MEDLINE-88233998; PubMed-2836808;
RA Hatt C., Ward M.E., Clarke I.N.;
RT "Analysis of the entire nucleotide sequence of the cryptic plasmid of
RL Chlamydia trachomatis serovar L1. Evidence for involvement in DNA
RT replication.";
RM Nucleic Acids Res. 16:4053-4067(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-G0/86 / SEROTYPE D: PLASMID-PCHL1;
RX MEDLINE-90301796; PubMed-2194229;
RA Comanducci M., Ricci S., Cevenini R., Ratti G.;
RT "Diversity of the Chlamydia trachomatis common plasmid in biovars
RL with different pathogenicity.";
RM Plasmid 23:149-154(1990).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-SEROTYPE B: PLASMID-PCTT1;
RX MEDLINE-88177106; PubMed-3444859;
RA Stripakash K.S., Macavoy E.S.;
RT "Characterization and sequence of a plasmid from the trachoma biovar
RL Plasmid 18:205-214(1987).
RN [5]
RP SEQUENCE OF 1-10.
RC STRAIN-L2/434/BU;
RA Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
RA Comanducci M., Christensen G., Birkelund S., Vretou E., Ratti G.,
RA Pallini V.;
RT Submitted (SEP-1994) to the SWISS-PROT data bank.
RN [6]
RP MISCELLANEOUS: PCR3-D IS REQUIRED FOR GROWTH WITHIN MAMMALIAN
CC CELLS.
CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF PLASMID PLGV440.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 126.
CC -1- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 38.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X07547; CAA30423.1; -
DR EMBL: X06707; CAA39894.1; -
DR EMBL: J03321; AAA91571.1; -
DR EMBL: M19487; AAB02589.1; ALT_FRAME.
DR PIR: S01922; S01922.
DR PIR: S00790; S00790.
DR PIR: S00791; S00791.
KM Plasmid.
FT INITMET 0 0
FT VARIANT 11 11 Q -> E (IN PLASMIDS PCHL1 AND PCTT1).
FT VARIANT 60 60 P -> S (IN PLASMIDS PCHL1 AND PCTT1).
FT VARIANT 85 85 D -> N (IN PLASMIDS PCTT1).
FT VARIANT 89 89 G -> D (IN PLASMIDS PCHL1 AND PCTT1).
FT VARIANT 108 108 T -> K (IN PLASMIDS PCHL1 AND PCTT1).
FT VARIANT 137 137 R -> S (IN PLASMIDS PCHL1 AND PCTT1).
FT VARIANT 188 188 K -> N (IN PLASMID PCTT1).
FT VARIANT 190 190 Y -> C (IN PLASMIDS PCHL1 AND PCTT1).

```

```

FT VARIANT 200 200 V -> I (IN PLASMID PCHL1).
FT VARIANT 209 209 R -> S (IN PLASMIDS PCHL1 AND PCTT1).
FT VARIANT 211 211 I -> T (IN PLASMIDS PCHL1 AND PCTT1).
SQ SEQUENCE 263 AA; 27791 MW; A3FF2469B1C89703 CRC64;

Query Match
Best local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STRVA 5
Db 39 STRVA 43

RESULT 30
HB2X_HUMAN STANDARD; PRT; 268 AA.
AC P05538;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DX BETA CHAIN PRECURSOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-87250501; PubMed-3036828;
RA Jonsson A.K., Hyldig-Nielsen J.J., Serenius B., Larhammar D.,
RA Andersson G., Joergensen F., Peterson P.A., Rask L.;
RT "Class II genes of the human major histocompatibility complex.
RT Comparisons of the DQ and DX alpha and beta genes.";
RL J. Biol. Chem. 262:8767-8777(1987).
RN [2]
RP SEQUENCE OF 38-125 FROM N.A.
RX MEDLINE-85216510; PubMed-3858830;
RA Okada K., Boss J.M., Prentice H., Spies T., Mengler R., Auftray C.,
RA Lillie J.W., Grossberger D., Strominger J.L.;
RT "Gene organization of DC and DX subregions of the human major
RT histocompatibility complex.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:3410-3414(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M29614; -; NOT_ANNOTATED_CDS.
DR EMBL: M29615; -; NOT_ANNOTATED_CDS.
DR EMBL: M1136; -; NOT_ANNOTATED_CDS.
DR PIR: D29312; D29312.
DR HSSP: P13760; 2SEB.
DR InterPro: IPR000353; -
DR InterPro: IPR000495; -
DR InterPro: IPR003006; -
DR Pfam: PF00969; MHC_II_beta; 1.
DR Pfam: PF00047; Iy; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KM MHC II; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 32
FT CHAIN 33 268 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN,
FT FT DX BETA CHAIN.
FT DOMAIN 33 126 EXTRACELLULAR BETA-1.
FT DOMAIN 127 229 EXTRACELLULAR BETA-2.
FT TRANSMEM 230 250
FT DOMAIN 251 268 CYTOPLASMIC TAIL.
FT DISULFID 47 110 BY SIMILARITY.
FT FT 148 204 BY SIMILARITY.

```

RT "Complete amino acid sequence of an HLA-DR antigen-like beta chain as  
 RT predicted from the nucleotide sequence: similarities with  
 RT Immunoglobulins and HLA-A, -B, and -C antigens."  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3687-3691(1982).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97419171; Pubmed=9271631;  
 RA Ellis M.C., Heltimer A.H., Ruddy D.A., Hansen S.L., Kronmal G.S.,  
 RA McClelland E., Quintana L., Drayna D.T., Aldrich M.S., Mlynar E.;  
 RT "HLA class II haplotype and sequence analysis support a role for DQ in  
 RT narcolepsy." 46:410-417(1997).  
 RL Immunogenetics 46:410-417(1997).  
 CC -I- MISCELLANEOUS: THE SEQUENCE SHOWN WAS PREVIOUSLY CALLED AN HLA-DR  
 CC BETA CHAIN BY REF.3 AUTHORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: K02405; AAA75521.1; -;  
 DR EMBL: U92032; AAB9191.1; -;  
 DR PIR: A02231; HLHDB.  
 DR HSSP: P13760; 2SEB.  
 DR InterPro: IPR000353; -;  
 DR InterPro: IPR000495; -;  
 DR InterPro: IPR003006; -;  
 DR Pfam: PF00969; MHC\_II\_beta; 1.  
 DR Pfam: PF00047; Ig; 1.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KW MHC II: Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 32  
 FT CHAIN 33 261  
 FT FT  
 FT DOMAIN 33 126 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN,  
 FT DOMAIN 127 220 DQ(1) BETA CHAIN.  
 FT DOMAIN 221 230 EXTRACELLULAR BETA-1.  
 FT TRANSMEM 231 251 EXTRACELLULAR BETA-2.  
 FT DOMAIN 252 261 CONNECTING PEPTIDE.  
 FT CYTOPLASMIC TAIL.  
 FT DISULFID 47 111 BY SIMILARITY.  
 FT DISULFID 149 205 BY SIMILARITY.  
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 167 167 D -> G (IN DOB1\*0202).  
 FT /FTID=VAR\_008115.  
 FT CONFLICT 69 69 I -> V (IN REF. 2 AND 3).  
 FT FT  
 FT SEQUENCE 261 AA; 29733 MW; 8957ACD3E93DD105 CRC64;  
 SQ  
 Query Match 41.7%; Score 5; DB 1; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 1.le+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 STPVA 5  
 DB 26 STPVA 30

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84070816; Pubmed=6316358;  
 RA Larhammar D., Hyldig-Nielsen J.J., Servenius B., Andersson G.,  
 RA Rask L., Peterson P.A.;  
 RT "Exon-Intron organization and complete nucleotide sequence of a human  
 RT major histocompatibility antigen DC beta gene."  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:7313-7317(1983).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87250501; Pubmed=3036828;  
 RA Jonsson A.K., Hyldig-Nielsen J.J., Servenius B., Larhammar D.,  
 RA Andersson G., Joergensen F., Peterson P.A., Rask L.;  
 RT "Class II genes of the human major histocompatibility complex.  
 RT Comparisons of the DQ and DX alpha and beta genes."  
 RL J. Biol. Chem. 262:8767-8777(1987).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: K01499; AAA98746.1; -;  
 DR EMBL: M29613; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL: M29616; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: A02234; HLH01C.  
 DR PIR: C29312; C29312.  
 DR HSSP: P13760; 2SEB.  
 DR InterPro: IPR000353; -;  
 DR InterPro: IPR000495; -;  
 DR InterPro: IPR003006; -;  
 DR Pfam: PF00969; MHC\_II\_beta; 1.  
 DR Pfam: PF00047; Ig; 1.  
 DR PROSITE: PS00290; IG\_MHC; 1.  
 KW MHC II: Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 32  
 FT CHAIN 33 261  
 FT FT  
 FT DOMAIN 33 126 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN,  
 FT DOMAIN 127 220 DQ(3) BETA CHAIN.  
 FT DOMAIN 221 230 EXTRACELLULAR BETA-1.  
 FT TRANSMEM 231 251 EXTRACELLULAR BETA-2.  
 FT DOMAIN 252 261 CONNECTING PEPTIDE.  
 FT CYTOPLASMIC TAIL.  
 FT DISULFID 47 111 BY SIMILARITY.  
 FT DISULFID 149 205 BY SIMILARITY.  
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 261 AA; 29870 MW; 8C2CFE6C080D4AF1 CRC64;  
 SQ  
 Query Match 41.7%; Score 5; DB 1; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 1.le+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 STPVA 5  
 DB 26 STPVA 30

RESULT 28  
 HB24\_HUMAN STANDARD: PRT: 261 AA.  
 ID HB24\_HUMAN  
 AC P01920;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DQ(3) BETA CHAIN PRECURSOR  
 DE (CLONE II-102).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;

RESULT 29  
 GP3D\_CHLTR STANDARD: PRT: 263 AA.  
 ID GP3D\_CHLTR  
 AC P10557; P08783; P08784; Q46431;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE VIRULENCE PLASMID PROTEIN RGP3-D (PROTEINS P-6/P-7).  
 OS Chlamydia trachomatis.  
 OG Plasmid pG6V440, Plasmid pCHL1, and Plasmid pCTT1.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;

RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 29342 / M129;  
 RX MEDLINE-96177562; PubMed-8604303;  
 RA Hilbert H., Himmelreich R., Plagens H., Herrmann R.;  
 RT "Sequence analysis of 56 kb from the genome of the bacterium  
 MYCOPLASMA PNEUMONIAE comprising the dnaA region, the atp operon and a  
 RT cluster of ribosomal protein genes.";  
 RL Nucleic Acids Res. 24:628-639(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 29342 / M129;  
 RX MEDLINE-97105885; PubMed-8948633;  
 RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae.";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U34816; AAC43647.1; -;  
 DR EMBL: AE000017; AAB95803.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 250 AA; 28662 MW; 4378A2834BBB7877 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 250;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYVK 12  
 II III  
 Db 24 OEYVK 28

RESULT 26  
 CB24\_ARATH STANDARD; PRT; 251 AA.  
 ID CB24\_ARATH  
 AC P27521;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CHLOROPHYLL A-B BINDING PROTEIN 4 PRECURSOR (LHCII TYPE III CAB-4)  
 DE (LHCP).  
 GN CAB4 OR P1P2.20.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicaceae; Brassicaceae; Arabidopsis.  
 OX NCBI\_Taxid=3702;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA Zhang H., Hanley S., Goodman H.M.;  
 RT "Isolation, characterization and chromosomal location of a new cab  
 RT gene from Arabidopsis thaliana.";  
 RL Plant Physiol. 96:1387-1388(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Choise N., Robert C., Brotier P., Catolico L.,  
 RA Alliqueneave F., Saurin W., Weissenbach J., Mewes H.-W., Lemcke K.,  
 RA Mayer K.F.X., Queller F., Salanoubat M.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: THE LIGHT-HARVESTING COMPLEX (LHC) FUNCTIONS AS A LIGHT  
 CC RECEPTOR. IT CAPTURES & DELIVERS EXCITATION ENERGY TO PHOTOSYSTEMS  
 CC WITH WHICH IT IS CLOSELY ASSOCIATED. THE N-TERMINUS OF THE PROTEIN  
 CC EXTENDS INTO THE STROMA WHERE IT IS INVOLVED WITH ADHESION OF

CC GRANAL MEMBRANES AND PHOTOREGULATED BY REVERSIBLE PHOSPHORYLATION  
 CC OF ITS THREONINE RESIDUES; BOTH ARE BELIEVED TO MEDIATE THE  
 CC DISTRIBUTION OF EXCITATION ENERGY BETWEEN PHOTOSYSTEMS I AND II.  
 CC -1- SUBUNIT: THE LHC COMPLEX CONSISTS OF CHLOROPHYLLS (A & B) AND  
 CC CHLOROPHYLL A-B BINDING PROTEINS.  
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M63931; AAA32760.1; -;  
 DR EMBL: AL132955; CAB61973.1; -;  
 DR InterPro: IPR001344; -;  
 DR Pfam: PF00504; chloroa\_b-bind. 1.  
 KW Chlorophyll; Photosynthesis; Photosystem I; Photosystem II;  
 KW Thylakoid membrane; Chloroplast; Transit peptide; Multigene family;  
 KW Transmembrane; Phosphorylation.  
 FT TRANSIT 1 ?  
 FT CHAIN 251 ?  
 FT TRANSMEM 211 227  
 FT POTENTIAL.  
 SQ SEQUENCE 251 AA; 27733 MW; DC8B390C2AED9D22 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APPOE 9  
 II III  
 Db 196 APPOE 200

RESULT 27  
 HB21\_HUMAN STANDARD; PRT; 261 AA.  
 ID HB21\_HUMAN  
 AC P01918; P01917;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DQ(1) BETA CHAIN PRECURSOR  
 DE (DC-3 BETA CHAIN).  
 GN HLA-DQB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-84298108; PubMed-6206493;  
 RA Boss J.M., Strominger J.L.;  
 RT "Cloning and sequence analysis of the human major histocompatibility  
 RT complex gene DC-3 beta.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5199-5203(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-84031733; PubMed-6415003;  
 RA Larhammar D., Andersson G., Andersson M., Bill P., Boehme J.,  
 RA Claesson L., Denaro M., Emmoth E., Gustafsson K., Hammarling U.,  
 RA Heldin E., Hyldig-Nielsen J.J., Lind P., Schenning L., Servenius B.,  
 RA Widmark E., Rask L., Peterson P.A.;  
 RT "Molecular analysis of human class II transplantation antigens and  
 RT their genes.";  
 RL Hum. Immunol. 8:95-103(1983).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-82247917; PubMed-6954511;  
 RA Larhammar D., Schenning L., Gustafsson K., Wiman K., Claesson L.,  
 RA Rask L., Peterson P.A.;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PROEV 10  
|||||

DB 118 PROEV 122

RESULT 23  
SOD3\_PLEBO STANDARD: PRT: 239 AA.

AC P50060:  
01-OCT-1996 (Rel. 34, Created)  
01-OCT-1996 (Rel. 34, Last sequence update)  
30-MAY-2000 (Rel. 39, Last annotation update)  
DE SUPEROXIDE DISMUTASE [MN] 3 PRECURSOR (EC 1.15.1.1) (FRAGMENT).  
GN SOD3.  
OS Plectonema boryanum.  
OC Bacteria; Cyanobacteria; Oscillatoriales; Plectonema.  
OX NCBI\_TaxID=1184;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-UTEX 485;  
RX MEDLINE=95164530; PubMed=7860607;  
RA Campbell W.S., Laudenbach D.E.;  
RT "Characterization of four superoxide dismutase genes from a filamentous cyanobacterium."  
RL J. Bacteriol. 177:964-972(1995).  
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.  
CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) -> O(2) + H(2)O(2).  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: 017611; AAA69953.1; -.  
CC DR HSSP: P09214; IMAAG.  
CC DR InterPro: IPR001189; -.  
CC DR Pfam: PF00081; sodef; 1.  
CC DR PROSITE: PS00088; SOD\_MN; 1.  
CC KM Oxidoreductase; Manganese; Multigene family; Signal.  
CC FT SIGNAL 1  
CC FT CHAIN 1  
CC FT METAL 31 30 POTENTIAL.  
CC FT METAL 57 57 SUPEROXIDE DISMUTASE [MN] 3.  
CC FT METAL 112 112 MANGANESE(BY SIMILARITY).  
CC FT METAL 195 195 MANGANESE(BY SIMILARITY).  
CC FT METAL 199 199 MANGANESE(BY SIMILARITY).  
CC FT METAL 199 199 MANGANESE(BY SIMILARITY).  
CC SQ SEQUENCE 239 AA; 26854 MW; 242CB3C9C61D100 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRPVA 5  
|||||

DB 18 SRPVA 22

RESULT 24  
Y984\_CAMJE STANDARD: PRT: 246 AA.  
AC P43491; Q9PNV5;  
DT 01-NOV-1995 (Rel. 32, Created)  
DR 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL PROTEIN CJ0984.  
GN CJ0984.  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
OC Campylobacter.  
OX NCBI\_TaxID=197;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 43431;  
RX MEDLINE=95247673; PubMed=7730270;  
RA Hani E.K., Chan V.L.;  
RT "Expression and characterization of Campylobacter jejuni benzoylglycine amidohydrolase (Hppuricase) gene in Escherichia coli."  
RL J. Bacteriol. 177:2396-2402(1995).  
RN (2)  
RP SEQUENCE FROM N.A.  
RC STRAIN-NCIT 11168;  
RX MEDLINE=20150912; PubMed=10688204;  
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C., Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S., Jagsis K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W., Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M., Whitehead S., Barrall B.G.;  
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences."  
RL Nature 403:665-668(2000).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: 236940; CAA85397.1; -.  
CC DR EMBL: AL139076; CAB73240.1; -.  
CC KM Hypothetical protein.  
CC FT CONFLICT 16 16 V -> T (IN REF. 1).  
CC FT CONFLICT 21 21 N -> D (IN REF. 1).  
CC FT CONFLICT 42 42 E -> G (IN REF. 1).  
CC FT CONFLICT 51 51 Q -> K (IN REF. 1).  
CC FT CONFLICT 118 118 I -> V (IN REF. 1).  
CC FT CONFLICT 160 160 E -> K (IN REF. 1).  
CC FT CONFLICT 212 212 V -> I (IN REF. 1).  
CC SQ SEQUENCE 246 AA; 28820 MW; 378876C1703D70CF CRC64;

Query Match 41.7%; Score 5; DB 1; Length 246;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APTOE 9  
|||||

DB 75 APTOE 79

RESULT 25  
Y887\_MYCPN STANDARD: PRT: 250 AA.  
AC Q50315;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL PROTEIN MPN687 (K05\_ORF250).  
GN MPN687 OR MP155.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmatidae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
RN (1)



Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,  
 Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
 Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 Ralback H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
 McDonald L., Arlrich P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 Venter J.C.;  
 RT Complete genome sequence of *Treponema pallidum*, the syphilis  
 agent.  
 RT *Science* 281:375-388(1998).  
 CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 5-PHOSPHATE - D-XYLULOSE 5-  
 PHOSPHATE.  
 CC -1- SIMILARITY: BELONGS TO THE RIBULOSE-PHOSPHATE 3-EPIMERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U97573; AAC08057.1; -  
 DR EMBL: AE001263; AAC65902.1; -  
 DR TIGR: TP0945; -  
 DR InterPro: IPR000056; -  
 DR Pfam: PFO0834; Ribul\_P\_3.epim.1.  
 DR PROSITE: PS01085; RIBUL\_P\_3-EPIMER\_1; 1.  
 DR PROSITE: PS01086; RIBUL\_P\_3-EPIMER\_2; 1.  
 KW Isomerase: Carbohydrate metabolism.  
 KW SEQUENCE 218 AA; 23514 MW; 0066263301FA2FCA CRC64;  
 SQ

Query Match 41.7%; Score 5; DB 1; Length 218;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVA 5  
 Db 120 STPVA 124

RESULT 21  
 TRKA\_PYRHO  
 ID TRKA\_PYRHO STANDARD; PRT; 228 AA.  
 AC 057719;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE TRK SYSTEM POTASSIUM UPTAKE PROTEIN TRKA HOMOLOG.  
 GN TRKA OR PH1884.  
 OS *Pyrococcus horikoshii*.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; *Pyrococcus*.  
 OX NCBI\_TaxID=53953;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OT3.  
 RX MEDLINE=98344137; PubMed=9679194.  
 RA Kawarabayashi Y., Sawada M., Horikawa H., Hatakeyama Y., Hino Y.,  
 Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 Funahashi T., Tanaka T., Kudoh Y., Yamazaki Y., Kushioka N., Oguchi A.,  
 Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 Masuuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3.";  
 RL DNA Res. 5:55-76(1998).  
 CC -1- FUNCTION: PART OF A POTASSIUM TRANSPORT SYSTEM (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE TRKA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP000007; BAA3111.1; -  
 KW Transport; Potassium transport; NAD.  
 FT NP\_BIND 2. 31 NAD (POTENTIAL).  
 SQ SEQUENCE 228 AA; 25323 MW; F64E2D14FE3FFD6 CRC64;  
 SQ

Query Match 41.7%; Score 5; DB 1; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVK 12  
 Db 215 QEVK 219

RESULT 22  
 KLR2\_CAVPO  
 ID KLR2\_CAVPO STANDARD; PRT; 239 AA.  
 AC P12323;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GLANDULAR KALLIKREIN, PROSTATIC (EC 3.4.21.35) (TISSUE KALLIKREIN)  
 DE (PROSTATE ESTERASE).  
 OS *Cavia porcellus* (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; *Cavia*.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=8800549; PubMed=3307909;  
 RX Dunbar J.C., Bradshaw R.A.;  
 RT "Amino acid sequence of guinea pig prostate kallikrein.";  
 RL *Biochemistry* 26:3471-3478(1987).  
 CC -1- FUNCTION: GLANDULAR KALLIKREIN CLEAVE MET-LYS AND ARG-SER BONDS  
 IN KININOGEN TO RELEASE LYS-BRADYKININ.  
 CC -1- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE OF ARG-1-XAA BONDS IN  
 SMALL MOLECULE SUBSTRATES. HIGHLY SELECTIVE ACTION TO RELEASE  
 KALLIDIN (LYSYL-BRADYKININ) FROM KININOGEN INVOLVES HYDROLYSIS OF  
 MET-1-XAA OR LEU-1-XAA.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 TRYPSIN FAMILY. KALLIKREIN SUBFAMILY.  
 CC -----  
 DR PIR: A27207; A27207.  
 DR HSP: P00757; 1SGF.  
 DR MEROPS: S01.160; -  
 DR InterPro: IPR001254; -  
 DR InterPro: IPR001314; -  
 DR Pfam: PF00089; TRYPSIN\_1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PROSITE: PS00134; TRYPSIN\_HIS\_1.  
 DR PROSITE: PS00135; TRYPSIN\_SER\_1.  
 KW Hydrolyase: Serine protease; Kininogenase; Glycoprotein.  
 FT ACT\_SITE 41  
 FT ACT\_SITE 96  
 FT ACT\_SITE 191  
 FT ACT\_SITE 191  
 FT CARBOHYD 78  
 FT CARBOHYD 169  
 FT CARBOHYD 169  
 FT DISULFID 7  
 FT DISULFID 151  
 FT DISULFID 26  
 FT DISULFID 42  
 FT DISULFID 128  
 FT DISULFID 197  
 FT DISULFID 162  
 FT DISULFID 176  
 FT DISULFID 187  
 FT DISULFID 212  
 FT VARIANT 50  
 FT VARIANT 50  
 SQ SEQUENCE 239 AA; 25989 MW; 56DC81BC10D49A64 CRC64;  
 SQ

Query Match 41.7%; Score 5; DB 1; Length 239;  
 Best Local Similarity 100.0%; Pred. No. 99;

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X13065; CAA31470.1; -  
 DR EMBL: X13065; CAA31469.1; -  
 DR PIR: S04827; S04827.  
 DR PIR: S04826; S04826.  
 KM Early protein.  
 FT CHAIN 1 209 PROTEIN 40A.  
 FT CHAIN 115 209 PROTEIN 40B.  
 SQ SEQUENCE 209 AA; 22974 MW; DD399BAA02E7643D CRC64;

Query Match 41.7%; Score 5; DB 1; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 QEVK 11  
 |||||  
 DB 59 QEVK 63

RESULT 18  
 V3P\_ADEA1 STANDARD; PRT: 217 AA.  
 ID V3P\_ADEA1  
 AC P19416; Q64818; (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DE 01-MAR-1992 (Rel. 21, Last annotation update)  
 DE 33 KDA PHOSPHOPROTEIN.  
 OS Human adenovirus type 41.  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 OX NCBI\_TaxID=10524;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TAK;  
 RA MEDLINE=90272433; PubMed=2349115;  
 RA Slemenda S.B., Pleniazek N.J., Velarde J. Jr., Pleniazek D.,  
 RA Luftig R.B.;  
 RT "Nucleotide sequence of the region coding for 100K and 33K proteins  
 RT of human enteric adenovirus type 41 (Tak).";  
 RL Nucleic Acids Res. 18:3069-3069(1990).  
 RN 12  
 RP SEQUENCE OF 203-217 FROM N.A.  
 RC STRAIN-TAK;  
 RX MEDLINE=89345113; PubMed=2762136;  
 RA Pleniazek N.J., Velarde J. Jr., Pleniazek D., Luftig R.B.;  
 RT "Nucleotide sequence of human enteric adenovirus type 41 hexon-  
 RT associated protein VIII precursor (PVIH) including the early region  
 RT E3 promoter.";  
 RL Nucleic Acids Res. 17:5398-5398(1989).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X52532; CAB38632.1; -  
 DR EMBL: X15137; CAA33236.1; -  
 DR PIR: S04851; S04851.  
 DR PIR: S10212; S10212.  
 KW late protein; Phosphorylation.  
 SQ SEQUENCE 217 AA; 24740 MW; B6ECDA6B3C24EC3E CRC64;

Query Match 41.7%; Score 5; DB 1; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 91;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 3 PVAPT 7  
 |||||  
 DB 118 PVAPT 122

RESULT 19  
 Y667\_SCHPO STANDARD; PRT: 217 AA.  
 ID Y667\_SCHPO  
 AC Q10319;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE HYPOTHETICAL 24.9 KDA PROTEIN C1768.07 IN CHROMOSOME I.  
 GN SPAC1768.07.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972;  
 RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO HUMAN AFG AND ENL AND YEAST TRANSCRIPTION  
 CC INITIATION FACTOR TFIIF SMALL SUBUNIT (TFIIF/ANCI).  
 CC -1- SIMILARITY: STRONG, TO YEAST YNL107W.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: 269795; CAA93690.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 217 AA; 24946 MW; A274C2AFB0620D7F CRC64;

Query Match 41.7%; Score 5; DB 1; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVK 12  
 |||||  
 DB 191 QEVK 195

RESULT 20  
 RPE\_TREPA STANDARD; PRT: 218 AA.  
 ID RPE\_TREPA  
 AC Q66107;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE RIBULOSE-5-PHOSPHATE 3-EPIMERASE (EC 5.1.3.1) (PENTOSE-5-PHOSPHATE 3-  
 DE EPIMERASE) (PPE) (R5P3E).  
 GN RPE OR TP0945.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=160;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Shevchenko D.V., Atkins D.R., Radolf J.D.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NICHOLS;  
 RX MEDLINE=98332770; PubMed=9665876;

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U39696; AAC71377.1; -.  
 DR TIGR: MG159; -.  
 DR InterPro: IPR001854; -.  
 DR Pfam: PF00831; Ribosomal\_L29; 1.  
 DR PROSITE: PS00579; RIBOSOMAL\_L29; 1.  
 KW Ribosomal protein.  
 SO SEQUENCE 200 AA; 23258 MW; 5BC115E52AA5EA92 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 200;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
 |||||  
 Db 170 OEYK 174

RESULT 15  
 ID IMMU\_BPSPB STANDARD; PRT; 201 AA.  
 AC P06650;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 01-MAR-1989 (Rel. 10, Last annotation update)  
 DE IMMUNITY PROTEIN.  
 GN D.  
 OS Bacteriophage SP-beta.  
 OC Viruses: dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;  
 OC Lambda phage group.  
 NC NCBL\_TaxID=10722;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C2;  
 RX MEDLINE=86304188; PubMed=3091583;  
 RA McLaughlin J.R., Wong H.C., Ting Y.E., van Arsdell J.N., Chang S.;  
 RT "Control of lysogeny and immunity of Bacillus subtilis temperate  
 RT bacteriophage SP beta by its d gene.";  
 RL J. Bacteriol. 167:952-959(1986).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M13821; AAA8544.1; -.  
 DR PIR: A24499; IMBPSB.  
 SO SEQUENCE 201 AA; 22087 MW; 25C0694D288FDAB CRC64;

Query Match 41.7%; Score 5; DB 1; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PROEV 10  
 |||||  
 Db 156 PROEV 160

RESULT 16  
 ID VADL\_TREPA STANDARD; PRT; 206 AA.  
 AC O83443;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE V-TYPE ATP SYNTHASE SUBUNIT D 1 (EC 3.6.1.34) (V-TYPE ATPASE SUBUNIT D  
 DE 1).  
 GN ATPD1 OR TP0428.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBL\_TaxID=160;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-NICHOLS;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
 RA McDonald L., Artach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 RT spirochete.";  
 RL Science 281:375-388(1998).  
 CC -I- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON  
 CC GRADIENT ACROSS THE MEMBRANE.  
 CC -I- SIMILARITY: BELONGS TO THE V-ATPASE D SUBUNIT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AE001220; AAC65414.1; -.  
 DR TIGR: TP0428; -.  
 DR InterPro: IPR002699; -.  
 DR Pfam: PF01813; ATP-synt\_D; 1.  
 KW Hydrolase; ATP synthesis; Hydrogen ion transport.  
 SO SEQUENCE 206 AA; 23803 MW; BC33C36FC5795483 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 206;  
 Best Local Similarity 100.0%; Pred. No. 87;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
 |||||  
 Db 80 OEYK 84

RESULT 17  
 ID VG40\_BPPH8 STANDARD; PRT; 209 AA.  
 AC P14817; P14818;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE PROTEINS 40A/40B.  
 GN 40.  
 OS Bacteriophage phi-80.  
 OC Viruses: dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;  
 OC Lambda phage group.  
 NC NCBL\_TaxID=10713;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89011978; PubMed=3172225;  
 RA Ogawa T., Ogawa H., Tomizawa J.;  
 RT "Organization of the early region of bacteriophage phi 80. Genes and  
 RT proteins.";  
 RL J. Mol. Biol. 202:537-550(1988).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

RA Eki T., Naitou M., Hagihara H., Ozawa M., Sasagawa S.-I.,  
 RA Sasagawa M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.,  
 RT "Analysis of a 36.2 kb DNA sequence including the right telomere of  
 RT chromosome VI from *Saccharomyces cerevisiae*.";  
 RL Yeast 12:149-167(1996).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL  
 CC INNER MEMBRANE (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D50617; BAA09284.1; -;  
 DR EMBL: D44597; BAA08011.1; -;  
 DR SGD: S0001941; YER045W.  
 DR InterPro: IPR001993; -;  
 DR Pfam: PF00153; mito\_carr; 1.  
 DR PROSITE: PS00215; MITOCH\_CARRIER; 1.  
 DR Hypothetical protein; Mitochondrion; Inner membrane; Repeat;  
 KW Transmembrane; Transport.  
 SO SEQUENCE 178 AA; 19901 MW; D57433A7D1AA4FE5 CRC64;

Query Match 41.7%; Score 5; DB 1; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 STPVA 5  
 Db 27 STPVA 31

RESULT 13  
 ID\_EFPL\_CHLPPN STANDARD: PRT; 185 AA.  
 AC Q92300; Q92076; (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE ELONGATION FACTOR P 1 (EF-P 1).  
 GN EFPI OR CEN0184 OR CP0584.  
 OS *Chlamydia pneumoniae* (*Chlamydia* pneumoniae).  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=83558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CHL029;  
 RA MEDLINE=9920606; PubMed=10192388;  
 RA Kalman S., Mitchell W., Marathe R., Lammell C., Fan J., Hyman R.W.,  
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
 RT "Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.";  
 RL Nat. Genet. 21:385-389(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AR39;  
 RA MEDLINE=20150255; PubMed=10684935;  
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
 RA White O., Hickey E.K., Peterson J., Umayam L.A., Utterback T.,  
 RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,  
 RA Bowman C., Dodson R., Gwinn M., Nelson W., Deboy R., Kolonay J.,  
 RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;  
 RT "Genome sequences of *Chlamydia trachomatis* Mopn and *Chlamydia*  
 RT *pneumoniae* AR39.";  
 RL Nucleic Acids Res. 28:1397-1406(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J138;  
 RA MEDLINE=20330349; PubMed=10871362;

RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
 RA Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;  
 RT "Comparison of whole genome sequences of *Chlamydia pneumoniae* J138  
 RT from Japan and CHL029 from USA.";  
 RL Nucleic Acids Res. 28:2311-2314(2000).  
 CC -1- FUNCTION: INVOLVED IN PEPTIDE BOND SYNTHESIS. STIMULATE EFFICIENT  
 CC TRANSLATION AND PEPTIDE-BOND SYNTHESIS ON NATIVE OR RECONSTITUTED  
 CC 70S RIBOSOMES IN VITRO. PROBABLY FUNCTIONS INDIRECTLY BY ALTERING  
 CC THE AFFINITY OF THE RIBOSOME FOR AMINOACYL-TRNA, THUS INCREASING  
 CC THEIR REACTIVITY AS ACCEPTORS FOR PEPTIDYL TRANSFERASE  
 CC (BY SIMILARITY).  
 CC -1- PATHWAY: PROTEIN BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE ELONGATION FACTOR P FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE001604; AAD18337.1; -;  
 DR EMBL: AE002217; AAF38402.1; -;  
 DR EMBL: AP002545; BAA98394.1; -;  
 DR TIGR: CP0584; -;  
 DR InterPro: IPR001059; -;  
 DR Pfam: PF01132; EFP; 1.  
 DR PROSITE: PS01275; EFP; 1.  
 KW Protein biosynthesis; Elongation factor.  
 SO SEQUENCE 185 AA; 20673 MW; 6C6432CFEB70354A CRC64;

Query Match 41.7%; Score 5; DB 1; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 TOEVR 11  
 Db 57 TOEVR 61

RESULT 14  
 ID\_RL29\_MYCGE STANDARD: PRT; 200 AA.  
 AC P47405;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE 50S RIBOSOMAL PROTEIN L29.  
 GN RPL29 OR RPL29 OR MG159.  
 OS *Mycoplasma genitalium*.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RA MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,  
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Bolt K.F., Hu P.-C., Lueker T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of *Mycoplasma genitalium*.";  
 RL Science 270:397-403(1995).  
 CC -1- SIMILARITY: BELONGS TO THE L29P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83300036; PubMed=6412231;
RA Emorine L., Dreher K.L., Kindt T.J., Max E.E.;
RT "Rabbit immunoglobulin kappa genes: structure of a germline b4
RT allele J-C locus and evidence for several b4-related sequences in
RT the rabbit genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:5709-5713(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=82060334; PubMed=6795636;
RA Heidmann O., Auffray C., Cazenave P.-A., Rougeon F.;
RT "Nucleotide sequence of constant and 3' untranslated regions of a
RT kappa immunoglobulin light chain mRNA of a homozygous b4 rabbit.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:5802-5806(1981).
RN [3]
RP SEQUENCE.
RX MEDLINE=75133568; PubMed=1091650;
RA Chen K.C.S., Kindt T.J., Krause R.M.;
RT "Primary structure of the L chain from a rabbit homogeneous antibody
RT to streptococcal carbohydrate. II. Sequence determination of peptides
RT from tryptic and peptic digests.";
RL J. Biol. Chem. 250:3289-3296(1975).
CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM ANTIBODY TO THE
CC SPECIFIC CARBOHYDRATE OF GROUP C STREPTOCOCCI & WAS ISOLATED FROM
CC THE SERUM OF A SINGLE RABBIT.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: X00231; CAA25051.1; -.
DR PIR: A02122; K4RB.
DR InterPro: IPR000495; -.
DR InterPro: IPR003006; -.
DR Pfam: PF00047; 19; 1.
DR PROSITE: PS00290; IG_MHC; FALSE_NEG.
KM Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 26 85
FT DISULFID 103 103
FT CONFLICT 58 58 N -> D (in REF. 3).
FT SEQUENCE 103 AA; 11043 MW; 5FC5ACCB8B60E68DB CRC64;
-----
Query Match 41.7%; Score 5; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 PVAPT 7
ID 11111
AC 2 PVAPT 6
Db

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83012203; PubMed=6214696;
RA Priest H., Kamp D., Kahmann R., Breuer B., Dellus H.;
RT "Nucleotide sequence of the immunity region of bacteriophage Mu.";
RL Mol. Gen. Genet. 186:315-321(1982).
RN [2]
RP SEQUENCE FROM N.A.
RA Priest H., Brauer B., Schmidt C., Kamp D.;
RT "Sequence of the left end of Mu.";
RL (in) Symonds N., Tousseint A., van de Putte P., Howe M.M. (eds.);
RL Phage Mu, pp.277-296, Cold Spring Harbor Laboratory Press,
RL New York (1987).
RN [3]
RP SEQUENCE FROM N.A.
RA Morgan G., Hatfull G., Hendrix R.;
RT "Genome of bacteriophage Mu and comparison with the Haemophilus
RT influenzae Mu-like prophage Flumu.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: V01464; CAA24711.1; -.
DR EMBL: M64097; AAA32376.1; -.
DR EMBL: AF083977; AAF01132.1; -.
DR PIR: S07291; S07291.
DR HSSP: P07636; ITNS.
KM Transcription regulation; Repressor; DNA-binding
SQ SEQUENCE 174 AA; 19213 MW; 11A4B3F5B85E49C5 CRC64;
-----
Query Match 41.7%; Score 5; DB 1; Length 174;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 8 OEYK 12
ID 11111
Db 168 OEYK 172

```

```

RESULT 11
RPL1_BPMU
ID RPL1_BPMU STANDARD; PRT; 174 AA.
AC P06019;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 01-OCT-2000 (Rel. 05, Last annotation update)
DE REPRESSOR PROTEIN CI.
GN CI OR 1.
OS Bacteriophage Mu.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.
OX NCBI_TaxID=10677;

```

```

RESULT 12
YFL5_YEAST
ID YFL5_YEAST STANDARD; PRT; 178 AA.
AC P43617;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE PUTATIVE MITOCHONDRIAL CARRIER YFR045W.
GN YFR045W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagihara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RX MEDLINE=96287652; PubMed=8686379;

```



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: A6000594; AAD07880.1; -  
 DR TIGR: HP0830; -  
 DR InterPro: IPR000120; -  
 DR Pfam: PF01425; Amidase; 1.  
 DR PROSITE: PS00571; AMIDASES; 1.  
 KW Protein biosynthesis; Ligase.  
 SO SEQUENCE 453 AA; 49652 MW; A35814B32F1AE13A CRC64;

Query Match  
 Best Local Similarity 50.0%; Score 6; DB 1; Length 453;  
 Pred. No. 15;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7  
 Db 373 TPVAPT 378

RESULT 6  
 ABPL\_SACEX STANDARD; PRT; 617 AA.  
 ID ABPL\_SACEX  
 AC p38479;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-FEB-1995 (Rel. 31, Last annotation update)  
 DE ACTIN BINDING PROTEIN.  
 GN ABPL  
 OS Saccharomyces exiguus (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=34358;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SRRAIN-CHS 379;  
 RX MEDLINE=94154001; PubMed=8110838;  
 RA Lange U., Steiner S., Grollig F., Wagner G., Philippsen P.;  
 RT Cloning and sequencing of a gene coding for an actin binding protein  
 RT of Saccharomyces exiguus; -  
 RL Biochim. Biophys. Acta 1217:214-218(1994).  
 CC -1- FUNCTION: MAY BE INVOLVED IN THE SPATIAL ORGANIZATION OF CELL.  
 CC SURFACE GROWTH. AN OVERPRODUCTION OF ABPL CAUSES THE ASSEMBLY OF  
 CC THE CORTICAL ACTIN SKELETON AT INAPPROPRIATE SITES ON THE CELL  
 CC SURFACE, RESULTING IN DELocalIZED SURFACE GROWTH (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CORTICAL CYTOSKELETON.  
 CC -1- SIMILARITY: TO SIGNAL TRANSDUCTION PROTEINS AND MYOSIN I.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: X73977; CAA52156.1; -  
 DR PIR: S42719; S42719.  
 DR HSSP: P19174; 1HSO.  
 DR InterPro: IPR001452; -  
 DR InterPro: IPR002108; -  
 DR Pfam: PF00018; SH3; 1.  
 DR Pfam: PF00241; cofilin.ADF; 1.  
 DR PRINTS: PR00452; SH3DOMAIN.  
 DR PROSITE: PSS0002; SH3; 1.  
 KW Cytoskeleton; Actin-binding; SH3 domain; Repeat.  
 FT DOMAIN 86 97 ACTIN-BINDING (POTENTIAL).

FT DOMAIN 557 617 SH3.  
 FT DOMAIN 486 501 POLY-GLU.  
 FT DOMAIN 523 529 POLY-GLU.  
 FT DOMAIN 202 600 3 X 10 AA APPROXIMATE REPEATS.  
 FT REPEAT 202 211 1-1.  
 FT REPEAT 444 453 1-2.  
 FT REPEAT 591 600 1-3.  
 FT DOMAIN 495 538 2 X 16 AA REPEATS OF E(7)-A-D-A-P-S-L-D-S-R.  
 FT REPEAT 495 510 2-1.  
 FT REPEAT 523 538 2-2.  
 FT REPEAT 523 538 2-2.  
 SO SEQUENCE 617 AA; 68420 MW; 540A785191B65F85 CRC64;

Query Match  
 Best Local Similarity 50.0%; Score 6; DB 1; Length 617;  
 Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12  
 Db 412 TOEVKK 417

RESULT 7  
 ENV\_FOAMV STANDARD; PRT; 985 AA.  
 ID ENV\_FOAMV  
 AC p14351;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE ENV POLYPROTEIN (COAT POLYPROTEIN).  
 GN ENV.  
 OS Human spumaretrovirus (Foamy virus).  
 OC Viruses; Retrovirdae; Spumavirus.  
 OX NCBI\_TaxID=11963;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88004420; PubMed=2820721;  
 RA Fluegel R.M., Retzlaff A., Maurer B., Darai G.;  
 RT Nucleotide sequence analysis of the env gene and its flanking  
 RT regions of the human spumaretrovirus reveals two novel genes; -  
 RL EMBO J. 6:2077-2084(1987).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: X05591; CAA29086.1; -  
 DR EMBL: M54978; AAA46123.1; -  
 DR PIR: C29685; VCLJSP.  
 KW Coat protein; Transmembrane; Polyprotein; Glycoprotein.  
 FT TRANSMEM 64 84 I (POTENTIAL).  
 FT TRANSMEM 958 978 II (POTENTIAL).  
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 282 282 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 342 342 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 524 524 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 553 553 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 805 805 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 830 830 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 985 AA; 113494 MW; 60F588E8A931BEOAF CRC64;

```
RESULT 3
VGG_BPPHK STANDARD: PRT: 187 AA.
AC 038042:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE MAJOR SPIKE PROTEIN (G PROTEIN) (GPC).
GN G.
OS Bacteriophage phi-K.
OC Viruses; ssDNA viruses; Microviridae; Microvirus.
ON NCBI_TaxID=10848;
RN [1]
RP SEQUENCE FROM N.A.
RL Kodaira K.-I., Tadokoro H., Suzuki K., Oki M., Manto S., Taketo A.;
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MAJOR SPIKE COMPONENT. INVOLVED IN THE ATTACHMENT TO
CC THE BACTERIAL HOST.
CC -1- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X60323; CAA42892.1; -.
DR HSSP; P03643; IAL0.
KW Cost protein.
SQ SEQUENCE 187 AA: 19569 MW: C5C5CECC891241701 CRC64:

Query Match
Best Local Similarity 50.0%; Score 6; DB 1; Length 187;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7
Db 25 TPVAPT 30

RESULT 4
GATA_HELPJ STANDARD: PRT: 453 AA.
ID GATA_HELPJ
AC 092L13:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE SUBUNIT A (EC 6.3.5.-) (GLU-ADT
DE SUBUNIT A).
GN GATA OR JHP0769.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
ON NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RL MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carnel G.,
RA Tumilimo P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang O., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RA "Genomic sequence comparison of two unrelated isolates of the human
RA gastric pathogen Helicobacter pylori."
RT Nature 397:176-180(1999).
RL -1- FUNCTION: FURNISHES A MEANS FOR FORMATION OF CORRECTLY CHARGED
CC GUN-TRNA(GLN) THROUGH THE TRANSAMIDATION OF MISACYLATED GLU-
CC TRNA(GLN) IN ORGANISMS WHICH LACK GLUTAMINYL-TRNA SYNTHETASE. THE
```

```
CC REACTION TAKES PLACE IN THE PRESENCE OF GLUTAMINE AND ATP THROUGH
CC AN ACTIVATED GAMMA-PHOSPHO-GLU-TRNA(GLN) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMYL-TRNA(GLN) + L-GLUTAMINE - ADP
CC + PHOSPHATE + L-GLUTAMINYL-TRNA(GLN) + L-GLUTAMATE.
CC -1- SUBUNIT: HETEROTRIMER OF A, B AND C SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AE001507; AAD06348.1; -.
DR InterPro; IPR000120; -.
DR Pfam; PF01425; Amidase; 1.
DR PROSITE; PS00571; AMIDASES; 1.
KW Protein biosynthesis; Ligase.
SQ SEQUENCE 453 AA: 49749 MW: FD97731532E77397 CRC64:

Query Match
Best Local Similarity 50.0%; Score 6; DB 1; Length 453;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7
Db 373 TPVAPT 378

RESULT 5
GATA_HELPY STANDARD: PRT: 453 AA.
ID GATA_HELPY
AC P56114:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE SUBUNIT A (EC 6.3.5.-) (GLU-ADT
DE SUBUNIT A).
GN GATA OR HP0830.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
ON NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RL STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Keriavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khaliq H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wattey L., Mallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RA "The complete genome sequence of the gastric pathogen Helicobacter
RA pylori."
RL Nature 388:539-547(1997).
CC -1- FUNCTION: FURNISHES A MEANS FOR FORMATION OF CORRECTLY CHARGED
CC GUN-TRNA(GLN) THROUGH THE TRANSAMIDATION OF MISACYLATED GLU-
CC TRNA(GLN) IN ORGANISMS WHICH LACK GLUTAMINYL-TRNA SYNTHETASE. THE
CC REACTION TAKES PLACE IN THE PRESENCE OF GLUTAMINE AND ATP THROUGH
CC AN ACTIVATED GAMMA-PHOSPHO-GLU-TRNA(GLN) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMYL-TRNA(GLN) + L-GLUTAMINE - ADP
CC + PHOSPHATE + L-GLUTAMINYL-TRNA(GLN) + L-GLUTAMATE.
CC -1- SUBUNIT: HETEROTRIMER OF A, B AND C SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

















GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:43:01 ; Search time 20.38 Seconds  
(without alignments)  
20.170 Million cell updates/sec

Title: US-09-372-036-30  
Perfect score: 12  
Sequence: 1 STPVAPTOEVKK 12

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size : 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	12	100.0	P60_LISMO	P21171 listeria mo
2	6	50.0	PRSI_SOLRU	P17641 solanum tub
3	6	50.0	187 VCG_BPPHK	Q38042 bacterioph
4	6	50.0	453 1 GATTA_HELPJ	Q92413 helicobacte
5	6	50.0	453 1 GATTA_HELPJ	P56114 helicobacte
6	6	50.0	617 1 ABPI_SACEX	P38479 saccharomyc
7	6	50.0	985 1 ENV_FOAMV	P14351 human spuma
8	6	50.0	1133 1 ATX9_TETH	Q05050 tetrathymena
9	5	41.7	102 1 HCC2_CRYCO	Q01238 cryptocodi
10	5	41.7	103 1 KAC4_RABIT	P01840 oryctolagus
11	5	41.7	174 1 RPCI_BPMU	P06019 bacterioph
12	5	41.7	178 1 YEL5_YEAST	P43617 saccharomyc
13	5	41.7	185 1 EEP1_CHLPH	Q92900 chlamydia p
14	5	41.7	200 1 RL29_MYCGE	P47405 mycoplasma
15	5	41.7	201 1 IMMU_BPSDB	P06650 bacterioph
16	5	41.7	206 1 VAD1_TREPA	O83443 treponema p
17	5	41.7	209 1 VG40_BPPH8	P14817 bacterioph
18	5	41.7	217 1 V33P_ADEA1	P19416 human adeno
19	5	41.7	217 1 YD67_SCHPO	Q10319 schizosacch
20	5	41.7	218 1 RPE_TREPA	O66107 treponema p
21	5	41.7	228 1 TRKA_PYRHO	O57719 pyrococcus
22	5	41.7	239 1 KLU2_CAVPO	P13233 cavia porce
23	5	41.7	239 1 SOD3_PLEBO	P50060 plectonema
24	5	41.7	246 1 Y984_CAMEE	P45491 campylobact
25	5	41.7	250 1 YG87_MYCPN	Q50313 mycoplasma
26	5	41.7	251 1 CB24_ARATH	P27521 arabidopsis
27	5	41.7	261 1 HB21_HUMAN	P01920 homo sapien
28	5	41.7	261 1 HB24_HUMAN	P01920 homo sapien
29	5	41.7	263 1 GP3D_CHLTR	P10557 chlamydia t
30	5	41.7	268 1 HB2X_HUMAN	P05538 homo sapien
31	5	41.7	270 1 URED_SYNY3	P73047 synechocyst
32	5	41.7	296 1 YB27_YEAST	P38279 saccharomyc
33	5	41.7	305 1 LMD2_PHOLE	Q06878 photobacter

34	5	41.7	309 1	POLG_HCVH7	P27955 hepatitis c
35	5	41.7	309 1	URIC_SOYBN	P04670 glycine max
36	5	41.7	309 1	URID_SOYBN	P04104 glycine max
37	5	41.7	314 1	ACCO_MALDO	O00985 malus domes
38	5	41.7	321 1	SPER_NPVOP	O65328 oryza pseu
39	5	41.7	323 1	PF27_MOUSE	P52875 mus musculu
40	5	41.7	325 1	MODD_MYCBO	O30620 mycobacteri
41	5	41.7	325 1	MODD_MYCRU	O50906 mycobacteri
42	5	41.7	327 1	EMT2_MOUSE	O61672 m equillipra
43	5	41.7	327 1	YB56_XENLA	P21574 xenopus lae
44	5	41.7	346 1	TRPA_MAIZE	P42390 zea mays (m
45	5	41.7	354 1	YMA3_MYCBO	Q02279 mycobacteri
46	5	41.7	356 1	HIS8_ECOLI	P06986 escherichia
47	5	41.7	359 1	HIS8_SALTY	P10369 salmonella
48	5	41.7	360 1	WMT2_CAEEL	P34889 caenorhabdi
49	5	41.7	362 1	SRK2_SPOLA	P42688 spongillia l
50	5	41.7	366 1	RPOO_REOVD	P03526 reovirus (t
51	5	41.7	366 1	RPOO_REOVL	P12002 reovirus (t
52	5	41.7	366 1	YMB5_YEAST	P07940 reovirus (t
53	5	41.7	367 1	MTCL_CHV13	Q04806 saccharomyc
54	5	41.7	367 1	NTIM_BOVAN	P36216 chlorella v
55	5	41.7	367 1	NTIM_BOVAN	P19041 podospora a
56	5	41.7	377 1	TTL_BOVIN	P38584 bos taurus
57	5	41.7	379 1	TTL_PIG	P38160 sus scrofa
58	5	41.7	380 1	YMH7_CAEEL	P34474 caenorhabdi
59	5	41.7	386 1	YGV5_YEAST	P53334 saccharomyc
60	5	41.7	388 1	GAG_HV1W2	P05889 human immun
61	5	41.7	390 1	EMRA_ECOLI	P27303 escherichia
62	5	41.7	391 1	THAB_PAROL	Q91242 paratubercu
63	5	41.7	400 1	CYH2_HUMAN	O99418 homo sapien
64	5	41.7	400 1	CYH2_MOUSE	P97695 mus musculu
65	5	41.7	415 1	YLN2_CAEEL	Q18964 caenorhabdi
66	5	41.7	428 1	SBK_HUMAN	P98077 homo sapien
67	5	41.7	432 1	RRA_XENLA	P24781 xenopus lae
68	5	41.7	442 1	TIG_CHLTR	O84713 chlamydia t
69	5	41.7	445 1	GLMM_HELPJ	Q95m22 helicobacte
70	5	41.7	445 1	GLMM_HELPJ	P25177 helicobacte
71	5	41.7	452 1	YDQ4_SCHPO	O14197 schizosacch
72	5	41.7	453 1	YDQ4_SCHPO	P34656 caenorhabdi
73	5	41.7	458 1	RRA_XENLA	P51126 xenopus lae
74	5	41.7	465 1	REBA_VIBCH	O07024 vibrio chol
75	5	41.7	468 1	PCOI_MOUSE	Q61398 mus musculu
76	5	41.7	468 1	PCOI_RAT	O08628 rattus norv
77	5	41.7	470 1	GLNA_FREDI	P33035 fremyella d
78	5	41.7	473 1	GLNA_SYNY3	P77961 synechocyst
79	5	41.7	473 1	SYE_AQUAE	O67271 aquiflex ae
80	5	41.7	473 1	YCLC_BACSU	P44005 bacillus su
81	5	41.7	474 1	VPQC_STRD7	O94697 streptomyc
82	5	41.7	474 1	VP61_NPVOP	O10270 oryza pseu
83	5	41.7	481 1	PE0_LISIN	O01836 listeria in
84	5	41.7	484 1	SGAT_ECOLI	P39301 escherichia
85	5	41.7	488 1	EXON_HSV6U	P24447 human herpe
86	5	41.7	488 1	EXON_HSV6Z	P52448 human herpe
87	5	41.7	488 1	PE24_RABIT	O26691 oryctolagus
88	5	41.7	490 1	SUCP_LEUME	O59495 leucostoc
89	5	41.7	498 1	ACHP_HUMAN	P30926 homo sapien
90	5	41.7	498 1	YCIK_ECOLI	P78061 escherichia
91	5	41.7	501 1	YH92_CAEEL	O22256 caenorhabdi
92	5	41.7	503 1	GAG_HV1JR	P20873 human immun
93	5	41.7	504 1	GAG_HV1MA	P04564 human immun
94	5	41.7	504 1	IRE5_HUMAN	O13568 homo sapien
95	5	41.7	509 1	CP77_BOVIN	O50185 bos taurus
96	5	41.7	509 1	PSBB_GUTHR	O78511 galliardia
97	5	41.7	509 1	VLI_HVY49	P36742 human papil
98	5	41.7	510 1	PURI_YEAST	P04046 saccharomyc
99	5	41.7	512 1	VLI_HPV24	P50790 human papil
100	5	41.7	514 1	VLI_HPV08	P06417 human papil
101	5	41.7	514 1	VLI_HPV47	P24242 human papil
102	5	41.7	515 1	VLI_HPV50	P50818 human papil
103	5	41.7	516 1	VLI_HPV05	P06917 human papil
104	5	41.7	516 1	VLI_HPV20	P50786 human papil
105	5	41.7	516 1	VLI_HPV36	P50812 human papil
106	5	41.7	517 1	VLI_HPV12	P36733 human papil



**This Page Blank (user)**

**THIS PAGE BLANK (USPTO)**

---

OY 2 TPVAP 6  
|||||  
Db 164 TPVAP 168

## RESULT 48

S07291  
repressor protein ci - phage Mu  
C:Species: phage Mu  
C>Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 08-Oct-1999  
C:Accession: S07291  
R:Pileups, H.; Kamp, D.; Kahmann, R.; Brauer, B.; Dellus, H.  
Mol. Gen. Genet. 186, 315-321, 1982  
A:Title: Nucleotide sequence of the immunity region of bacteriophage Mu.  
A:Reference number: S07291; MUID:83012203  
A:Accession: S07291  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-174 <PRI>  
A:Cross-references: EMBL:V01464; NID:915807; PIDN:CAA24711.1; PID:915808  
C:Keywords: DNA binding; transcription regulation

Query Match 41.7%; Score 5; DB 2; Length 174;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
|||||  
Db 168 OEYK 172

## RESULT 49

D86699  
hypothetical protein yfii [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 23-Mar-2001  
C:Accession: D86699  
R:Bojorlin, A.; Mucker, P.; Mauger, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. in press, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium.  
A:Reference number: A86625  
A:Accession: D86699  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-175 <SMO>  
A:Cross-references: GB:AE005176; NID:912723488; PIDN:AAK04694.1; GSPDB:GND0146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: yfii

Query Match 41.7%; Score 5; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
|||||  
Db 94 OEYK 98

## RESULT 50

T50849  
hypothetical protein 110K5.14 [imported] - sorghum (fragment)  
C:Species: Sorghum bicolor (sorghum)  
C>Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
C:Accession: T50849  
R:Tikhonov, A.P.; SanMiguel, P.J.; Nakajima, Y.; Gorenstein, N.M.; Bennetzen, J.L.; Ayre  
Proc. Natl. Acad. Sci. U.S.A. 96, 7409-7414, 1999  
A:Title: Colinearity and its exceptions in orthologous adh regions of maize and sorghum.  
A:Reference number: Z25261; MUID:99307423  
A:Accession: T50849  
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-175 <TIK>  
A:Cross-references: EMBL:AF124045; PIDN:AAD43043.1  
A:Experimental source: cultivar BTx623  
C:Genetics:  
A:Map position: J  
A>Note: 110K5.14

Query Match 41.7%; Score 5; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
|||||  
Db 68 TPVAP 72

Search completed: August 15, 2001, 12:34:44  
Job time: 109 sec

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: J33327  
 R:Madson, C.; Wamsley, P.; O'Brien, D.  
 A:Reference number: 733327  
 A:Submitted to the EMBL Data Library, July 1998  
 A:Description: The sequence of *Caenorhabditis elegans* cosmid C40D2.  
 A:Reference number: 221323  
 A:Accession: J33327  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Cross-references: EMBL:AF07530; PIDN:AAC26267.1; GSPDB:GN00020; CESP:C40D2.4  
 A:Residues: 1-147 <MAD>  
 A:Molecule type: DNA  
 A:Experimental source: strain Bristol N2; clone C40D2  
 A:Genetics:  
 A:Gene: CESP:C40D2.4  
 A:Map position: 2  
 A:Introns: 45/3; 58/1; 75/1; 102/3

Query Match 41.7%; Score 5; DB 2; Length 147;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APTOE 9  
 |||||  
 Db 53 APTOE 57

RESULT 44  
 T25638  
 hypothetical protein C46H11.5 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T25638  
 R:Miller, N.; Bradshaw, H.; Wamsley, P.  
 A:Submitted to the EMBL Data Library, February 1997  
 A:Description: The sequence of *C. elegans* cosmid C46H11.  
 A:Reference number: 220061  
 A:Accession: T25638  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-157 <MLI>  
 A:Cross-references: EMBL:U88314; PIDN:AAB42359.1; GSPDB:GN00019; CESP:C46H11.5  
 A:Experimental source: strain Bristol N2; clone C46H11  
 A:Genetics:  
 A:Gene: CESP:C46H11.5  
 A:Map position: 1  
 A:Introns: 12/1; 44/2

Query Match 41.7%; Score 5; DB 2; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APTOE 9  
 |||||  
 Db 86 APTOE 90

RESULT 45  
 C96568  
 hypothetical protein F6D8.6 [imported] - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: C96568  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Holt, D.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: C96568  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-161 <STO>  
 A:Cross-references: GB:AE05173; NID:95903065; PIDN:AAD55624.1; GSPDB:GN00141  
 A:Genetics:  
 A:Gene: F6D8.6  
 A:Map position: 1

Query Match 41.7%; Score 5; DB 2; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAP 7  
 |||||  
 Db 40 PVAP 44

RESULT 46  
 S61230  
 cytochrome-c biosynthesis heme-carrier protein cycJ - *Rhizobium leguminosarum*  
 C:Species: *Rhizobium leguminosarum*  
 C:Date: 18-Sep-1997 #sequence\_revision 18-Sep-1997 #text\_change 29-Sep-1999  
 C:Accession: S61230  
 R:Delgado, M.J.; Yeaman, K.H.; Wu, G.; Vargas, C.; Davies, A.; Poole, R.K.; Johnston, submitted to the EMBL Data Library, July 1995  
 A:Description: Characterization of the *cycHKL* genes involved in cytochrome c biogenesis  
 A:Reference number: S61229  
 A:Accession: S61230  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-165 <DEL>  
 A:Cross-references: EMBL:X89726; NID:9967064; PIDN:CAA61877.1; PID:9967066  
 A:Superfamily: Escherichia coli cytochrome-c biosynthesis heme-carrier protein come  
 C:Keywords: carrier protein; chromoprotein; heme; iron; metalloprotein  
 F124/Binding site: heme (His) (covalent) #status predicted

Query Match 41.7%; Score 5; DB 2; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
 |||||  
 Db 45 TPVAP 49

RESULT 47  
 S11770  
 ImaA protein - *Listeria monocytogenes*  
 C:Species: *Listeria monocytogenes*  
 C:Date: 03-May-1994 #sequence\_revision 20-Feb-1995 #text\_change 24-Nov-1999  
 C:Accession: S11770  
 R:Goehmann, S.; Leimister-Waechter, M.; Schiltz, E.; Goebel, W.; Chakraborty, T.  
 Mol. Microbiol. 4, 1091-1099, 1990  
 A:Title: Characterization of a *Listeria monocytogenes*-specific protein capable of inducing apoptosis  
 A:Reference number: S11770; MUID:91041717  
 A:Accession: S11770  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-170 <GOE>  
 A:Genetics:  
 A:Gene: ImaA  
 C:Superfamily: *Listeria monocytogenes* ImaA protein

Query Match 41.7%; Score 5; DB 2; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

A:Reference number: 225035  
A:Accession: T50083  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1133 <SAU>  
A:Cross-references: EMBL:AL13984; PIDN:CAB61215.1; GSPDB:GN00066; SPDB:SPAC1556.04c  
A:Experimental source: strain 972h(-); cosmid c1556  
C:Genetics:  
A:Gene: SPDB:SPAC1556.04c  
A:Map position: 1  
A:Introns: 47/1; 67/3; 95/2; 121/3  
C:Superfamily: cytidine deaminase

Query Match 41.7%; Score 5; DB 2; Length 133;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
|||||  
DB 11 OEYK 15

RESULT 39  
B60497  
MHC class II histocompatibility antigen RT1 D-I beta chain precursor - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 12-Aug-1996  
C:Accession: B60497  
R:Holowachuk, E.W.; Greer, M.K.  
Diabetes 38, 267-271, 1989  
A:title: Unaltered class II histocompatibility antigens and pathogenesis of IDDM in BB  
A:Reference number: A60497; MUID:89121214  
A:Accession: B60497  
A:Molecule type: mRNA  
A:Residues: 1-134 <HOL>  
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplanted;  
F:1-27/Domain: signal sequence #status predicted <SIG>

Query Match 41.7%; Score 5; DB 2; Length 134;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPYA 5  
|||||  
DB 21 STPYA 25

RESULT 40  
A81676  
conserved hypothetical protein TC0687 [imported] - Chlamydia muridarum (strain N199)  
C:Species: Chlamydia muridarum, Chlamydia trachomatis Mopn  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Sep-2000  
C:Accession: A81676  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwyn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150235  
A:Accession: A81676  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-134 <RET>  
A:Cross-references: GB:AE002337; GB:AE002160; NID:g7190714; PIDN:AAF39504.1; PID:g719071  
C:Genetics:  
A:Experimental source: strain N199 (Mopn)  
C:Superfamily: Escherichia coli dosage-dependent dnaX suppressor protein dxsa

Query Match 41.7%; Score 5; DB 2; Length 134;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
|||||  
DB 41 OEYK 45

RESULT 41  
E72580  
hypothetical protein APE1924 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
C:Accession: E72580  
R:Kawabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta-  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.  
DNA Res. 6, 83-101, 1999  
A:title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero  
A:Reference number: A72450; MUID:99310339  
A:Accession: E72580  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-141 <KAW>  
A:Cross-references: DDBJ:AP000062; NID:95105244; PIDN:BAA80930.1; PID:d1044716; PID:g  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE1924

Query Match 41.7%; Score 5; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 VAPPO 8  
|||||  
DB 66 VAPPO 70

RESULT 42  
G72384  
conserved hypothetical protein - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: G72384  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwyn, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,  
C.M.  
Nature 399, 323-329, 1999  
A:title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MUID:99287316  
A:Accession: G72384  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-147 <ARN>  
A:Cross-references: GB:AE001717; GB:AE000512; NID:g4980871; PIDN:AAD35456.1; PID:g498  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM0369

Query Match 41.7%; Score 5; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
|||||  
DB 127 OEYK 131

RESULT 43  
T33327  
hypothetical protein C40D2.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans

A:Accession: G72066  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-124 <ARN>  
 A:Cross-references: GB:AE001638, GB:AE001363; NID:94376819; PIDN:AA18674.1; PID:9437682  
 A:Experimental source: strain CWL029  
 C:Genetics:  
 A:Gene: dksA  
 C:Superfamily: Escherichia coli dosage-dependent dnaK suppressor protein dksA

Query Match 41.7%; Score 5; DB 2; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVKK 12  
 |||||  
 DB 31 QEVKK 35

## RESULT 34

T24876  
 hypothetical protein T13F2.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jun-2000  
 C:Accession: T24876  
 R:Swingburne, J.  
 submitted to the EMBL Data Library, October 1996  
 A:Reference number: Z19947  
 A:Accession: T24876  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-124 <MIL>  
 A:Cross-references: EMBL, Z81122; PIDN:CA803353.1; GSPDB:GN00022; CESP:T13F2.2  
 A:Experimental source: clone T13F2  
 C:Genetics:  
 A:Gene: CESP:T13F2.2  
 A:Map position: 4  
 A:Introns: 27/3; 61/3; 91/3  
 C:Superfamily: Caenorhabditis elegans hypothetical protein T13F2.2

Query Match 41.7%; Score 5; DB 2; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVKK 12  
 |||||  
 DB 39 QEVKK 43

## RESULT 35

I54454  
 lymphocyte antigen - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999  
 C:Accession: I54454  
 R:Hiratawa, A.; Seyfried, C.E.; Nepom, G.T.; Milner, E.C.  
 Immunogenetics 29, 186-190, 1989  
 A:Title: Sequence analysis of HLA class II domains: characterization of the DQw3 family  
 A:Reference number: I54454; MUID:89173071  
 A:Accession: I54454  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-126 <RES>  
 A:Cross-references: GB:M25325; NID:9619801; PIDN:AAA59675.1; PID:9619802  
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 41.7%; Score 5; DB 2; Length 126;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRPVA 5  
 |||||  
 DB 26 SRPVA 30

## RESULT 36

D86684  
 prophage p11 protein 41, tail component [imported] - Lactococcus lactis subsp. lactis  
 C:Species: Lactococcus lactis subsp. lactis  
 C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 23-Mar-2001  
 C:Accession: D86684  
 R:Boletín, A.; Winkler, P.; Mauger, S.; Jallón, O.; Malarne, K.; Weissenbach, J.; Eh  
 genome Res. in press, 2001  
 A:Title: The complete genome sequence of the lactic acid bacterium.  
 A:Reference number: AB6625

A:Accession: D86684  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-131 <STO>  
 A:Cross-references: GB:AE005176; NID:912723355; PIDN:AKK04574.1; GSPDB:GN00146  
 A:Experimental source: strain IL1403  
 C:Genetics:  
 A:Gene: p1141

Query Match 41.7%; Score 5; DB 2; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PTQEV 10  
 |||||  
 DB 3 PTQEV 7

## RESULT 37

B81600  
 conserved hypothetical protein CP0218 [imported] - Chlamydomonada pneumoniae (strain A  
 C:Species: Chlamydomonada pneumoniae, Chlamydia pneumoniae  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Sep-2000  
 C:Accession: B81600  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
 , C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe  
 Nuclear Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39  
 A:Reference number: AB1500; MUID:20150255  
 A:Accession: B81600  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-132 <REA>  
 A:Cross-references: GB:AE002183; GB:AE002161; NID:97189146; PIDN:ARF38087.1; PID:9718  
 A:Experimental source: strain AR39, HL cells  
 C:Genetics:  
 A:Gene: CP0218  
 C:Superfamily: Escherichia coli dosage-dependent dnaK suppressor protein dksA

Query Match 41.7%; Score 5; DB 2; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVKK 12  
 |||||  
 DB 39 QEVKK 43

## RESULT 38

T50083  
 probable cytidine deaminase [imported] - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 21-Jul-2000  
 C:Accession: T50083  
 R:Saunders, D.; Harris, D.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, November 1999

A:Residues: 1-114 <UE>  
 A:Cross-references: EMBL:AL049862; GSPDB:GND0061; ATSP:FI8B3.180  
 A:Experimental source: cultivar Columbia; BAC clone FI8B3  
 C:Genetics:  
 A:Gene: ATSP:FI8B3.180  
 A:Map position: 3  
 C:Superfamily: Arabidopsis hypothetical protein FI8B3.180

Query Match 41.7%; Score 5; DB 2; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVKK 12  
 |||||  
 DB 43 QEVKK 47

RESULT 29  
 B49094  
 methylmalonyl-CoA decarboxylase (EC 4.1.1.41) delta chain - Veillonella parvula  
 C:Species: Veillonella parvula  
 C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 05-Nov-1999  
 C:Accession: B49094  
 R:Huder, J.B.; Dimroth, P.  
 J. Biol. Chem. 268, 24564-24571, 1993  
 A:Title: Sequence of the sodium ion pump methylmalonyl-CoA decarboxylase from Veillonella  
 A:Reference number: A49094; MID:94043308  
 A:Accession: B49094  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-115 <HUD>  
 A:Cross-references: GB:L22208; NID:9415592; PIDN:AAC36821.1; PID:9415594  
 C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 41.7%; Score 5; DB 2; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
 |||||  
 DB 60 TPVAP 64

RESULT 30  
 A72293  
 conserved hypothetical protein - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: A72293  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
 C.M.  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
 A:Reference number: A72200; MID:99287316  
 A:Accession: A72293  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-117 <ARN>  
 A:Cross-references: GB:AE001770; GB:AE000512; NID:94981658; PIDN:AAD36200.1; PID:9498167  
 C:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: TM1124

Query Match 41.7%; Score 5; DB 2; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPYA 5  
 |||||

DB 30 STPYA 34

RESULT 31  
 B86557  
 Dnak suppressor [imported] - Chlamydia pneumoniae (strain J139)  
 C:Species: Chlamydia pneumoniae  
 C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
 C:Accession: B86557  
 R:Shiral, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;  
 Nucleic Acids Res. 28, 2311-2314, 2000  
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.  
 A:Reference number: A86491; MID:20330349  
 A:Accession: B86557  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-124 <STO>  
 A:Cross-references: GB:BA000008; NID:98978905; PIDN:BA98740.1; GSPDB:GND0142  
 A:Experimental source: strain J138  
 C:Genetics:  
 A:Gene: dksA  
 C:Superfamily: Escherichia coli dosage-dependent dnak suppressor protein dksA

Query Match 41.7%; Score 5; DB 2; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVKK 12  
 |||||  
 DB 31 QEVKK 35

RESULT 32  
 A71518  
 probable dnak suppressor - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C>Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 09-Jun-2000  
 C:Accession: A71518  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche  
 Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t  
 A:Reference number: A71570; MID:99000809  
 A:Accession: A71518  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-124 <ARN>  
 A:Cross-references: GB:AE001314; GB:AE001273; NID:93328833; PIDN:AAC68004.1; PID:9332  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: dksA  
 C:Superfamily: Escherichia coli dosage-dependent dnak suppressor protein dksA

Query Match 41.7%; Score 5; DB 2; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVKK 12  
 |||||  
 DB 31 QEVKK 35

RESULT 33  
 G72066  
 dnak suppressor - Chlamydia pneumoniae (strain CWL029)  
 C:Species: Chlamydia pneumoniae  
 C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jun-2000  
 C:Accession: G72066  
 R:Rahman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,  
 Nature Genet. 21, 385-389, 1999  
 A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
 A:Reference number: A72000; MID:99206606

Db 31 STPVA 35

## RESULT 24

B56581  
 major basic nuclear protein, variant HCC2 - dinoflagellate (Cryptocodinium cohnii) (fr  
 N:Alternate names: protein alpha chain p14  
 C:Species: Cryptocodinium cohnii  
 C:Date: 08-Sep-1995 #sequence\_revision 08-Sep-1995 #text\_change 21-Jul-2000  
 C:Accession: B56581; S14640; S14644  
 R:Salta-Rovira, M.; Geraud, M.L.; Caput, D.; Jacques, F.; Soyer-Gobillard, M.O.; Vermet,  
 Chromosoma 100, 510-518, 1991  
 A:Title: Molecular cloning and immunolocalization of two variants of the major basic nuc  
 A:Reference number: A56581; MUID:92111321  
 A:Accession: B56581  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-102 <SAL>  
 A:Cross-references: EMBL:X58444; NID:q17962; PIDN:CAA41350.1; PID:q17963  
 A>Note: sequence extracted from NCBI backbone (NCBIN:77122, NCBIPI:77123)  
 C:Keywords: DNA binding; nucleus

Query Match 41.7%; Score 5; DB 2; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEVRK 12  
 |||||  
 Db 43 OEVRK 47

## RESULT 25

K4RB  
 19 kappa-B4 chain C region - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 31-Jan-1981 #sequence\_revision 15-Oct-1982 #text\_change 16-Aug-1996  
 C:Accession: A93971; A93891; A92176; A02122  
 R:Emorline, L.; Dreher, K.; Kindt, T.J.; Max, E.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 5709-5713, 1983  
 A:Title: Rabbit immunoglobulin kappa genes: structure of a germline b4 allotype J-C locu  
 A:Reference number: A93971; MUID:83300036  
 A:Accession: A93971  
 A:Molecule type: DNA  
 A:Residues: 1-103 <EMO>  
 A>Note: the sequence was determined from the germline gene  
 R:Heidmann, O.; Auftray, C.; Cazeneuve, P.A.; Rougeon, F.  
 Proc. Natl. Acad. Sci. U.S.A. 78, 5802-5806, 1981  
 A:Title: Nucleotide sequence of constant and 3' untranslated regions of a kappa immunogl  
 A:Reference number: A93891; MUID:82060334  
 A:Accession: A93891  
 A:Molecule type: mRNA  
 A:Residues: 1-103 <HER>  
 R:Chen, K.C.S.; Kindt, T.J.; Krause, R.M.  
 J. Biol. Chem. 250, 3289-3296, 1975  
 A:Title: Primary structure of the L chain from a rabbit homogeneous antibody to streptoc  
 A:Reference number: A92176; MUID:75133568  
 A:Accession: A92176  
 A:Molecule type: protein  
 A:Residues: 1-57, 'D', '59-103 <CHR>  
 A>Note: this chain was obtained from antibody to the specific carbohydrate of group C st  
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: heterotetramer  
 F:19-87/Domain: immunoglobulin homology <IMM>

Query Match 41.7%; Score 5; DB 1; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7

Db 2 PVAPT 6  
 |||||

## RESULT 26

F53275  
 Ig kappa-1 chain C region b95 allotype - rabbit (fragment)  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 02-May-1994 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000  
 C:Accession: F53275  
 R:Ayadi, H.; Marche, P.N.; Cazeneuve, P.A.  
 Immunogenetics 34, 201-207, 1991  
 A:Title: Evolution of the rabbit immunoglobulin kappa chain genes.  
 A:Reference number: A53275; MUID:91372868  
 A:Accession: F53275  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-104 <AYA>  
 A>Note: sequence inconsistent with nucleotide translation  
 A:Note: sequence extracted from NCBI backbone (NCBIN:56069, NCBIPI:56170)  
 C:Superfamily: immunoglobulin C region; immunoglobulin homology  
 C:Keywords: heterotetramer; immunoglobulin  
 F:19-87/Domain: immunoglobulin homology <IMM>

Query Match 41.7%; Score 5; DB 2; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
 |||||  
 Db 2 PVAPT 6

## RESULT 27

114788  
 hypothetical protein DKFZP564A122.1 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T14788  
 R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, August 1999  
 A:Reference number: Z18178  
 A:Accession: T14788  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-112 <DUE>  
 A:Cross-references: EMBL:AL110269  
 A:Experimental source: fetal brain; clone DKFZP564A122  
 C:Genetics:  
 A>Note: DKFZP564A122.1

Query Match 41.7%; Score 5; DB 2; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PTOEV 10  
 |||||  
 Db 91 PTOEV 95

## RESULT 28

T08411  
 hypothetical protein F18b3.180 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 22-Oct-1999  
 C:Accession: T08411  
 R:Quetier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salan  
 submitted to the Protein Sequence Database, May 1999  
 A:Reference number: Z16409  
 A:Accession: T08411  
 A:Molecule type: DNA



T30302  
P-type ATPase - Tetrahymena thermophila  
C:Species: Tetrahymena thermophila  
C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 07-Dec-1999  
C:Accession: T30302  
R:Wang, S.; Takeyasu, K.  
Submitted to the EMBL Data Library, November 1995  
A:Description: The starvation-induced P-type ATPase in Tetrahymena thermophila.  
A:Reference number: 220815  
A:Accession: T30302  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1133 <MAN>  
A:Cross-references: EMBL:U41063; NID:g1545827; PID:g1545828; PIDN:AAB08071.1  
C:Genetics:  
A:Genetic code: SGCS  
A:Note: TPA9

Query Match 50.0%; Score 6; DB 2; Length 1133;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEYKK 12  
|||||  
DB 1099 TOEYKK 1104

RESULT 20  
E64664  
Outer membrane protein - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: E64664  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khatlek, H.G.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97354467  
A:Accession: E64664  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1230 <TOM>  
A:Cross-references: GB:AE000621; GB:AE000511; NID:g2314301; PIDN:AAD08205.1; PID:g231431

Query Match 50.0%; Score 6; DB 2; Length 1230;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEYKK 12  
|||||  
DB 173 TOEYKK 178

RESULT 21  
E75264  
Hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: E75264  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamthyan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zaleski, C.; M. Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: E75264  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-75 <MHI>

A:Cross-references: GB:AE002081; GB:AE000513; NID:g6460337; PIDN:AAF12064.1; PID:g646  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR2520  
A:Map position: 1

Query Match 41.7%; Score 5; DB 2; Length 75;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 VAPTO 8  
|||||  
DB 24 VAPTO 28

RESULT 22  
T29802  
Hypothetical protein C06E4.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T29802  
R:Du, Z.; Gattung, S.  
Submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid C06E4.  
A:Reference number: 220688  
A:Accession: T29802  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-82 <DUZ>  
A:Cross-references: EMBL:U41277; PIDN:AAA82477.1; CESP:C06E4.2  
C:Genetics:  
A:Gene: CESP:C06E4.2  
A:Introns: 31/1

Query Match 41.7%; Score 5; DB 2; Length 82;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 VAPTO 8  
|||||  
DB 40 VAPTO 44

RESULT 23  
T34429  
Hypothetical protein F55C7.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34429  
R:Du, Z.; Le, T.  
Submitted to the EMBL Data Library, November 1996  
A:Description: The sequence of C. elegans cosmid F55C7.  
A:Reference number: 221524  
A:Accession: T34429  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-88 <DUZ>  
A:Cross-references: EMBL:U80436; PIDN:AAC71107.1; GSPDB:GN00019; CESP:F55C7.2  
A:Experimental source: strain Bristol N2; clone F55C7  
C:Genetics:  
A:Gene: CESP:F55C7.2  
A:Map position: 1  
A:Introns: 18/3

Query Match 41.7%; Score 5; DB 2; Length 88;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPYA 5  
|||||

submitted to the EMBL Data Library, October 1997

A:Reference number: Z19673

A:Accession: T23087

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-460 <WIL>

A:Cross-references: EMBL:Z99942; PIDN:CAB17068.1; GSPDB:GN00028; CESP:HL3N06.2

A:Experimental source: clone HL3N06

C:Genetics:

A:Gene: CESP:HL3N06.2

A:Map position: X

A:Introns: 17/1; 294/1; 331/3; 364/1; 414/2; 432/2

Query Match 50.0%; Score 6; DB 2; Length 460;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7

Db 73 TPVAPT 78

RESULT 15

actin-binding protein ABP1 - *Saccharomyces exiguus*

C:Species: *Saccharomyces exiguus*

C>Date: 07-Sep-1994 #sequence\_revision 19-Apr-1996 #text\_change 13-Aug-1999

C:Accession: S42719

R:Langre, U.; Steiner, S.; Grollig, F.; Wagner, G.; Philippson, P.

Biochim. Biophys. Acta 1217, 214-218, 1994

A:Title: Cloning and sequencing of a gene coding for an actin binding protein of *Sacchar*

A:Reference number: S42719; MUID:94154001

A:Accession: S42719

A:Molecule type: DNA

A:Residues: 1-617 <LAN>

A:Cross-references: EMBL:X73977; NID:G433513; PIDN:CAA52156.1; PID:G433514

C:Superfamily: actin-binding protein ABP1; SH3 homology

C:Keywords: actin binding; cytoskeleton

F:564-613/Domain: SH3 homology <SH3>

Query Match 50.0%; Score 6; DB 2; Length 617;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12

Db 412 TOEVKK 417

RESULT 16

hypothetical protein YOR070c - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: hypothetical protein Q2931

C:Species: *Saccharomyces cerevisiae*

C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 05-Nov-1999

C:Accession: S66953

R:Bonin, C.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Dang, D.V.; Valens, M.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66929

A:Accession: S66953

A:Molecule type: DNA

A:Residues: 1-637 <BOH>

A:Cross-references: EMBL:Z74978; NID:G1420220; PID:E251984; PID:G1420221; GSPDB:GN00015;

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YOR070c

A:Map position: 15R

Query Match 50.0%; Score 6; DB 2; Length 637;

Best Local Similarity 100.0%; Pred. No. 47;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7

Db 181 TPVAPT 186

RESULT 17

HB4913

probable ATP-dependent RNA helicase [imported] - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: H84913

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon,

euss, D.; Neriman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487

A:Accession: H84913

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-748 <STO>

A:Cross-references: GB:AE002093; NID:G2275211; PIDN:AAB63833.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g47330

A:Map position: 2

Query Match 50.0%; Score 6; DB 2; Length 748;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPRQ 8

Db 725 PVAPRQ 730

RESULT 18

VCLJSP

env polyprotein - human foamy virus

N:Alternate names: coat polyprotein

C:Species: human foamy virus

A>Note: host *Homo sapiens* (man)

C>Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 16-Jul-1999

C:Accession: C29685

R:Fluegel, R.M.; Reithelm, A.; Maurer, B.; Darai, G.

EMBO J. 6, 2077-2084, 1987

A:Title: Nucleotide sequence analysis of the env gene and its flanking regions of the

A:Reference number: A91074; MUID:88004420

A:Accession: C29685

A:Molecule type: genomic RNA

A:Residues: 1-985 <FLU>

A:Cross-references: GB:X05591; GB:Y00070; NID:G61759; PIDN:CAA29086.1; PID:G61762

C:Genetics:

A:Gene: env

C:Superfamily: foamy virus env polyprotein

C:Keywords: coat protein; polyprotein; transmembrane protein

F:64-87/Domain: transmembrane #status predicted <TM1>

F:579-595/Domain: transmembrane #status predicted <TM2>

F:936-972/Domain: transmembrane #status predicted <TM3>

Query Match 50.0%; Score 6; DB 1; Length 985;

Best Local Similarity 100.0%; Pred. No. 69;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12

Db 181 TOEVKK 186

RESULT 19

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-293 <NH1>  
A:Cross-references: GB:AE001940; GB:AE000513; NID:g6458577; PIDN:AAF10447.1; PID:g645858  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0864  
A:Map position: 1

Query Match 50.0%; Score 6; DB 2; Length 293;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPVAPT 6  
|||||  
DB 227 TPVAPT 232

RESULT 10

T34992  
probable lipoprotein - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T34992  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1998  
A:Reference number: 221550  
A:Accession: T34992  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-447 <OLI>  
A:Cross-references: EMBL:AL031182; PIDN:CAA20169.1; GSPDB:GN00070; SCOEDB:SC4A2.17c  
C:Genetics:  
A:Gene: SCOEDB:SC4A2.17c

Query Match 50.0%; Score 6; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7  
|||||  
DB 63 TPVAPT 68

RESULT 11

F64623  
amidase - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: F64623  
R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalaf, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.; Nature 388, 539-547, 1997  
A:Authors: Mellin, E.; Hayes, W.S.; Bordovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97354467  
A:Accession: F64623  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-453 <ROM>  
A:Cross-references: GB:AE000594; GB:AE000511; NID:g2313957; PIDN:AAD07880.1; PID:g231396  
C:Superfamily: Indoleacetamide hydrolase

Query Match 50.0%; Score 6; DB 1; Length 453;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7

DB 373 TPVAPT 378  
|||||

RESULT 12

A71891  
glu-tRNA amidotransferase, chain A - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 22-Jun-1999  
C:Accession: A71891  
R:Lim, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.  
Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
A:Reference number: A71800; MUID:99120557  
A:Accession: A71891  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-453 <ARN>  
A:Cross-references: GB:AE001507; GB:AE001439; NID:g4155326; PIDN:AAD06348.1; PID:g415  
C:Genetics:  
A:Experimental source: strain J99  
C:Gene: gata  
C:Superfamily: Indoleacetamide hydrolase

Query Match 50.0%; Score 6; DB 2; Length 453;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7  
|||||  
DB 373 TPVAPT 378

RESULT 13

H85574  
hypothetical protein Z0891 [Imported] - Escherichia coli (strain O157:H7)  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
C:Accession: H85574  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
hiller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallantha, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: H85574  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-456 <STO>  
A:Cross-references: GB:AE005174; NID:g12513654; PIDN:AA655060.1; GSPDB:GN00145; UNCP:  
A:Experimental source: strain O157:H7, substrain EDL533  
C:Genetics:  
A:Gene: Z0891

Query Match 50.0%; Score 6; DB 2; Length 456;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 VAPTOE 9  
|||||  
DB 378 VAPTOE 383

RESULT 14

T23087  
hypothetical protein HJ3N06.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T23087  
R:Lennard, N.

Db 29 TOEVKK 34

RESULT 5

S35162

STH-21 protein - potato

C:Species: Solanum tuberosum (potato)

C:Date: 13-Jan-1995 #sequence\_revision 09-May-1997 #text\_change 20-Aug-1999

C:Accession: S35162; S11869

R:Maton, D.P.; Prescott, G.; Bertrand, C.; Camirand, A.; Brisson, N.

Plant Mol. Biol. 22, 279-291, 1993

A:Title: Identification of cis-acting elements involved in the regulation of the pathogen

A:Reference number: S35161; MUID:93283632

A:Accession: S35162

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-155 <MAN>

A:Cross-references: EMBL:M29042; NID:q169577; PIDN:AAA02829.1; PID:q169578

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1990

C:Genetics:

A:Gene: STH-21

A:Introns: 57/1

C:Superfamily: pathogenesis-related protein

Query Match

Best Local Similarity 50.0%; Score 6; DB 2; Length 155;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7

Db 12 TPVAPT 17

RESULT 6

JC4806

core protein G - phage phi-K

C:Species: Phage phi-K

C:Date: 15-Aug-1996 #sequence\_revision 15-Oct-1996 #text\_change 20-Jun-2000

C:Accession: JC4806; B04253; A04253

R:Kodaira, K.; Oki, M.; Kakikawa, M.; Kimoto, H.; Takeko, A.

J. Biochem. 119, 1062-1069, 1996

A:Title: The virion proteins encoded by bacteriophage phi-K and its host-range mutant ph

A:Reference number: JC4804; MUID:96424987

A:Accession: JC4806

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-187 <KOD>

A:Cross-references: EMBL:X60323; NID:q1478118; PIDN:CAA42892.1; PID:q1478127

R:Slins, J.; Capon, D.; Dressler, D.

J. Biol. Chem. 254, 12615-12628, 1979

A:Title: drug (primase)-dependent origins of DNA replication. Nucleotide sequences of th

A:Reference number: A92247; MUID:80049950

A:Accession: B04253

A:Molecule type: DNA

A:Residues: 165-187 <SIM>

C:Comment: This protein is one of the structural components of the bacteriophage capsid.

C:Genetics:

A:Gene: G

C:Superfamily: phage phi-X174 gene G protein

C:Keywords: capsid protein

Query Match

Best Local Similarity 50.0%; Score 6; DB 2; Length 187;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7

Db 25 TPVAPT 30

RESULT 7

C86847

transcription regulator Tena [imported] - Lactococcus lactis subsp. lactis (strain IL

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 31-Mar-2001

C:Accession: C86847

R:Boitlin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Eh

Genome Res. in press, 2001

A:Title: The complete genome sequence of the lactic acid bacterium.

A:Reference number: A86625

A:Accession: C86847

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-218 <STO>

A:Cross-references: GB:AF005176; NID:q12724803; PIDN:AAK05877.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: tena

C:Superfamily: transcription activator tena

Query Match

Best Local Similarity 50.0%; Score 6; DB 2; Length 218;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAPT 7

Db 103 TPVAPT 108

RESULT 8

F71504

hypothetical protein CT504 - Chlamydia trachomatis (serotype D, strain UM3/Cx)

C:Species: Chlamydia trachomatis

C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 18-Aug-2000

C:Accession: F71504

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche

Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t

A:Reference number: A71570; MUID:99000809

A:Accession: F71504

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-288 <ARN>

A:Cross-references: GB:AF001323; GB:AF001273; NID:q3328931; PIDN:AA68105.1; PID:q3332

A:Experimental source: serotype D, strain UM-3/Cx

C:Genetics:

A:Gene: CT504

C:Superfamily: conserved hypothetical protein TC0791

Query Match

Best Local Similarity 50.0%; Score 6; DB 2; Length 288;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVAP 6

Db 131 STPVAP 136

RESULT 9

F75466

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C:Accession: F75466

R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: F75466

979 4 33.3 128 2 H84223  
980 4 33.3 128 2 JN0727  
981 4 33.3 128 2 S75366  
982 4 33.3 128 2 T05635  
983 4 33.3 128 2 T32789  
984 4 33.3 129 2 A33548  
985 4 33.3 129 2 S63260  
986 4 33.3 129 2 S46393  
987 4 33.3 129 2 B71350  
988 4 33.3 129 2 S53806  
989 4 33.3 129 2 S37713  
990 4 33.3 129 2 F83764  
991 4 33.3 130 1 F64376  
992 4 33.3 130 1 C69883  
993 4 33.3 130 2 PC6025  
994 4 33.3 130 2 S08079  
995 4 33.3 130 2 F25733  
996 4 33.3 130 2 G70813  
997 4 33.3 130 2 S77271  
998 4 33.3 130 2 D70966  
999 4 33.3 130 2 T29887  
1000 4 33.3 130 2 F64348

## ALIGNMENTS

## RESULT 1

protein p60 precursor - Listeria monocytogenes  
N:Alternate names: Invasion-associated protein  
C:Species: Listeria monocytogenes  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 13-Sep-1996  
C:Accession: A41487 B41487  
R:Kochler, S.; Leimeister-Waechter, M.; Chakraborty, T.; Lottspeich, F.; Goebel, W.  
Interf. Immun. 58, 1943-1950, 1990  
A:Title: The gene coding for protein p60 of Listeria monocytogenes and its use as a spec  
A:Reference number: A41487; MUID:90256283  
A:Accession: A41487  
A:Molecule type: DNA  
A:Residues: 1484 <KOE>  
A:Cross-references: GB:X52268  
A:Accession: B41487  
A:Molecule type: protein  
A:Residues: 28-49 <KO2>  
C:Genetics:  
A:Gene: lbp  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-484/Product: protein 60 #status predicted <MAT>

Query Match 100.0%; Score 12; DB 2; Length 484;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STVPAPTOEVKK 12  
Db 146 STVPAPTOEVKK 157

## RESULT 2

probable tonB transport protein Cj1630 [Imported] - Campylobacter jejuni (strain NCTC 11  
C:Species: Campylobacter jejuni  
C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 31-Mar-2000  
C:Accession: D81259  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barre  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: A81250; MUID:20150912  
A:Accession: D81259  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-227 <PAR>  
A:Cross-references: GB:AL1139079; GB:AL111168; NID:g96968971; PIDN:CAB73618.1; PID:g966  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: tonB2; Cj1630

Query Match 58.3%; Score 7; DB 2; Length 227;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APTOEVKK 11  
Db 105 APTOEVKK 111

## RESULT 3

F83231  
hypothetical protein PA3307 [Imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83231  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A:Reference number: A82850; MUID:20437337  
A:Accession: F83231  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-102 <STO>  
A:Cross-references: GB:AE004753; GB:AE004091; NID:g9949433; PIDN:AA06695.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3307

Query Match 50.0%; Score 6; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 9.1;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPTQ 8  
Db 27 PVAPTQ 32

## RESULT 4

T37186  
probable regulatory protein - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T37186  
R:Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, July 1999  
A:Reference number: Z21598  
A:Accession: T37186  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-116 <SEE>  
A:Cross-references: EMBL:AL096823; PIDN:CAB46971.1; GSPDB:GN00070; SCOEDB:SC011.16  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC011.16  
C:Superfamily: Streptomyces ambifaciens replication activator protein pra

Query Match 50.0%; Score 6; DB 2; Length 116;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVKK 12

833	4	33.3	115	2	E45213	Ca2+-transporting	906	4	33.3	122	2	PC4279	anti-SS-A/Ro 60k p
834	4	33.3	115	2	PH1560	Ig heavy chain V r	907	4	33.3	122	2	PC4280	anti-SS-A/Ro 60k p
835	4	33.3	115	2	PH1557	Ig heavy chain V r	908	4	33.3	122	2	PH1426	Ig heavy chain V r
836	4	33.3	115	2	G72568	hypothetical prote	909	4	33.3	122	2	PH0958	Ig heavy chain V r
837	4	33.3	115	2	E70702	hypothetical prote	910	4	33.3	122	2	S36271	Ig heavy chain V r
838	4	33.3	115	2	E84684	hypothetical prote	911	4	33.3	122	2	C49590	Ig heavy chain V r
839	4	33.3	115	2	E86809	hypothetical prote	912	4	33.3	122	2	B49590	Ig heavy chain V r
840	4	33.3	115	2	D84501	hypothetical prote	913	4	33.3	122	2	T22096	hypothetical prote
841	4	33.3	115	2	E84512	probable MYB famll	914	4	33.3	123	1	RKSPS	ribulose-bisphosph
842	4	33.3	115	2	PC2230	hobo-like transpos	915	4	33.3	123	2	D33548	Ig heavy chain V-I
843	4	33.3	116	1	B46279	guanylin precursor	916	4	33.3	123	2	PH1413	Ig heavy chain V r
844	4	33.3	116	2	S31698	Ig heavy chain pre	917	4	33.3	123	2	PH1423	Ig heavy chain V r
845	4	33.3	116	2	PH0959	Ig heavy chain V r	918	4	33.3	123	2	S38492	Ig heavy chain - h
846	4	33.3	116	2	S36261	Ig heavy chain V r	919	4	33.3	123	2	C36006	Ig heavy chain V r
847	4	33.3	116	2	S31667	Ig heavy chain V r	920	4	33.3	123	2	S44108	Ig heavy chain V-D
848	4	33.3	116	2	PM0618	ribosomal protein	921	4	33.3	123	2	E64418	conserved hypotnet
849	4	33.3	116	2	UC5188	sporulation-specif	922	4	33.3	123	2	G71143	hypothetical prote
850	4	33.3	116	2	PC4228	N-acetylglucosamin	923	4	33.3	123	2	H75059	hypothetical prote
851	4	33.3	116	2	G70831	hypothetical prote	924	4	33.3	123	2	B75546	hypothetical prote
852	4	33.3	116	2	G70726	replication activa	925	4	33.3	123	2	G71189	hypothetical prote
853	4	33.3	116	2	T33195	hypothetical prote	926	4	33.3	123	2	C71423	hypothetical prote
854	4	33.3	116	2	H72489	hypothetical prote	927	4	33.3	123	2	A05125	hypothetical prote
855	4	33.3	117	1	GLHUEU	Ig heavy chain V-I	928	4	33.3	123	2	T20279	hypothetical prote
856	4	33.3	117	1	HVHU35	Ig heavy chain pre	929	4	33.3	124	1	T1BHB	trypsin inhibitor
857	4	33.3	117	1	HVHU35	Ig heavy chain pre	930	4	33.3	124	1	M1HUML	Ig heavy chain V-I
858	4	33.3	117	2	T02872	Ig heavy chain pre	931	4	33.3	124	1	M1HUST	Ig heavy chain V-I
859	4	33.3	117	2	S19669	probable lipid tra	932	4	33.3	124	1	MNIHMS	nonstructural prot
860	4	33.3	117	2	S19670	Ig heavy chain V r	933	4	33.3	124	2	S19665	Ig heavy chain V r
861	4	33.3	117	2	A28846	Ig heavy chain pre	934	4	33.3	124	2	E72403	Ig heavy chain V r
862	4	33.3	117	2	PT0371	Ig gamma chain pre	935	4	33.3	124	2	B70472	conserved hypotnet
863	4	33.3	117	2	S21668	Ig kappa chain V r	936	4	33.3	124	2	H72263	conserved hypotnet
864	4	33.3	117	2	S31680	Ig heavy chain V r	937	4	33.3	124	2	G75116	hypothetical prote
865	4	33.3	117	2	S18551	Ig heavy chain V r	938	4	33.3	124	2	B82196	ReB1 protein VC14
866	4	33.3	117	2	S18553	Ig heavy chain V r	939	4	33.3	124	2	T36292	hypothetical prote
867	4	33.3	117	2	S18552	Ig heavy chain V r	940	4	33.3	125	1	S1PSDR	steroid Delta-15om
868	4	33.3	117	2	B82614	hypothetical prote	941	4	33.3	125	1	HVHUMO	Ig heavy chain V-I
869	4	33.3	117	2	H86640	ribonuclease P (EC	942	4	33.3	125	2	PH1410	Ig heavy chain V r
870	4	33.3	117	2	E75537	conserved hypotnet	943	4	33.3	125	2	PH0957	Ig heavy chain V r
871	4	33.3	118	2	PH1666	Ig heavy chain V r	944	4	33.3	125	2	A64502	hypothetical prote
872	4	33.3	118	2	S36265	Ig heavy chain V r	945	4	33.3	125	2	S78697	probable export pr
873	4	33.3	118	2	S12440	Ig lambda chain (M	946	4	33.3	125	2	H86722	hypothetical prote
874	4	33.3	118	2	S12441	Ig lambda chain (K	947	4	33.3	125	2	T38537	probable single-st
875	4	33.3	118	2	S12442	Ig lambda chain (K	948	4	33.3	126	2	PH1417	Ig heavy chain V r
876	4	33.3	118	2	T35747	hypothetical prote	949	4	33.3	126	2	PH1418	Ig heavy chain V r
877	4	33.3	118	2	T07905	low-carbon dioxide	950	4	33.3	126	2	PH1424	Ig heavy chain V r
878	4	33.3	118	2	S34346	hypothetical prote	951	4	33.3	126	2	PH1412	Ig heavy chain V r
879	4	33.3	119	2	S08077	Ig kappa chain pre	952	4	33.3	126	2	B33548	Ig heavy chain V-I
880	4	33.3	119	2	JN0295	Ig heavy chain V-D	953	4	33.3	126	2	T24181	hypothetical prote
881	4	33.3	119	2	PH0961	Ig heavy chain V r	954	4	33.3	126	2	A56657	PIEMP2/MESA protei
882	4	33.3	119	2	S44106	Ig heavy chain V-D	955	4	33.3	126	2	G83571	conserved hypotnet
883	4	33.3	119	2	S58445	lipoprotein D - Sa	956	4	33.3	126	2	H75433	hypothetical prote
884	4	33.3	119	2	T00151	hypothetical prote	957	4	33.3	127	1	C64140	probable glycyL ra
885	4	33.3	119	2	E83674	hypothetical prote	958	4	33.3	127	2	PH1415	Ig heavy chain V r
886	4	33.3	119	2	G86843	hypothetical prote	959	4	33.3	127	2	PH1415	Ig heavy chain V r
887	4	33.3	119	2	D64002	hypothetical prote	960	4	33.3	127	2	PH1421	Ig heavy chain V r
888	4	33.3	119	2	E72526	hypothetical prote	961	4	33.3	127	2	PH1411	Ig heavy chain V r
889	4	33.3	120	2	S21667	Ig kappa chain V r	962	4	33.3	127	2	PH1420	Ig heavy chain V r
890	4	33.3	120	2	S21666	Ig kappa chain V r	963	4	33.3	127	2	PH0955	Ig heavy chain V r
891	4	33.3	120	2	PH0962	Ig heavy chain V r	964	4	33.3	127	2	S34014	Ig heavy chain V r
892	4	33.3	120	2	T29774	hypothetical prote	965	4	33.3	127	2	S00653	chlrophyll a/b-b1
893	4	33.3	120	2	D71039	hypothetical prote	966	4	33.3	127	2	F82197	RseB2 protein VC14
894	4	33.3	120	2	T23978	hypothetical prote	967	4	33.3	127	2	G85091	arabinogalactan-pr
895	4	33.3	120	2	C85849	unknown protein en	968	4	33.3	127	2	G83870	inner spore coat p
896	4	33.3	120	2	A12539	hypothetical prote	969	4	33.3	128	2	I69024	MHC sex-limited pr
897	4	33.3	121	2	E70313	histidine triad-li	970	4	33.3	128	2	S16685	Ig heavy chain V r
898	4	33.3	121	2	A29678	T-cell receptor ga	971	4	33.3	128	2	PH0952	Ig heavy chain V r
899	4	33.3	121	2	S20783	Ig heavy chain V r	972	4	33.3	128	2	S76468	hypothetical prote
900	4	33.3	121	2	A49590	Ig heavy chain V r	973	4	33.3	128	2	T22276	hypothetical prote
901	4	33.3	121	2	UC4622	ribosomal phosphop	974	4	33.3	128	2	D75575	5-carboxymethyl-2-
902	4	33.3	121	2	D75089	hypothetical prote	975	4	33.3	128	2	S76955	hypothetical prote
903	4	33.3	121	2	T32888	hypothetical prote	976	4	33.3	128	2	D72481	hypothetical prote
904	4	33.3	121	2	S74021	hypothetical prote	977	4	33.3	128	2	H70435	hypothetical prote
905	4	33.3	121	2	S75660	hypothetical prote	978	4	33.3	128	2	T15911	hypothetical prote

687	4	33.3	101	2	S12424	Ig heavy chain V r	760	4	33.3	106	2	S64649	RBL2 protein - yea
688	4	33.3	101	2	S12428	Ig heavy chain V r	761	4	33.3	107	2	S40290	protein-tyrosine-p
689	4	33.3	101	2	S12431	Ig heavy chain V r	762	4	33.3	107	2	S67118	hypothetical prote
690	4	33.3	101	2	H65009	hypothetical prote	763	4	33.3	107	2	T45621	hypothetical prote
691	4	33.3	101	2	B72079	hypothetical prote	764	4	33.3	107	2	B64845	hypothetical prote
692	4	33.3	101	2	F86544	hypothetical prote	765	4	33.3	107	2	F83182	hypothetical prote
693	4	33.3	101	2	S76550	hypothetical prote	766	4	33.3	107	2	E83040	DNA-binding protei
694	4	33.3	101	2	T19560	hypothetical prote	767	4	33.3	107	2	G81080	hypothetical prote
695	4	33.3	102	1	W7MLDP	E7 protein - deer	768	4	33.3	108	2	PH1664	Ig heavy chain V r
696	4	33.3	102	1	W7MLDP	E7 protein - Europ	769	4	33.3	108	2	C49056	T-cell receptor al
697	4	33.3	102	2	PH1232	Ig heavy chain V r	770	4	33.3	108	2	SS1687	protein-tyrosine-p
698	4	33.3	102	2	PH1234	Ig heavy chain V r	771	4	33.3	108	2	TS1146	ring-box protein 1
699	4	33.3	102	2	PH1235	Ig heavy chain V r	772	4	33.3	108	2	A86675	hypothetical prote
700	4	33.3	102	2	PH1237	Ig heavy chain V r	773	4	33.3	108	2	S61627	hypothetical prote
701	4	33.3	102	2	PH1238	Ig heavy chain V r	774	4	33.3	108	2	D72652	hypothetical prote
702	4	33.3	102	2	PH1239	Ig heavy chain V r	775	4	33.3	108	2	A72516	hypothetical prote
703	4	33.3	102	2	PH1240	Ig heavy chain V r	776	4	33.3	108	2	T26681	hypothetical prote
704	4	33.3	102	2	PH1243	Ig heavy chain V r	777	4	33.3	108	2	F84479	En/Spm-like transp
705	4	33.3	102	2	PH1248	Ig heavy chain V r	778	4	33.3	108	2	A72704	hypothetical prote
706	4	33.3	102	2	PH1262	Ig heavy chain V r	779	4	33.3	109	2	S16437	ribulose-bisphosph
707	4	33.3	102	2	PH1264	Ig heavy chain V r	780	4	33.3	109	2	PH1672	Ig heavy chain V r
708	4	33.3	102	2	PH1268	Ig heavy chain V r	781	4	33.3	109	2	PH1668	Ig heavy chain V r
709	4	33.3	102	2	PH1276	Ig heavy chain V r	782	4	33.3	109	2	PH1671	Ig heavy chain V r
710	4	33.3	102	2	PH1280	Ig heavy chain V r	783	4	33.3	109	2	D71279	hypothetical prote
711	4	33.3	102	2	PH1281	Ig heavy chain V r	784	4	33.3	109	2	F84010	hypothetical prote
712	4	33.3	102	2	PH1244	Ig heavy chain V r	785	4	33.3	109	2	T34639	hypothetical prote
713	4	33.3	102	2	PH1247	Ig heavy chain V r	786	4	33.3	109	2	TS1864	probable heat-shoc
714	4	33.3	102	2	PH1249	Ig heavy chain V r	787	4	33.3	110	1	A37414	parvalbumin - mous
715	4	33.3	102	2	PH1252	Ig heavy chain V r	788	4	33.3	110	2	PH1669	Ig heavy chain V r
716	4	33.3	102	2	PH1254	Ig heavy chain V r	789	4	33.3	110	2	PH1670	Ig heavy chain V r
717	4	33.3	102	2	PH1258	Ig heavy chain V r	790	4	33.3	110	2	G70305	ribosomal protein
718	4	33.3	102	2	PH1259	Ig heavy chain V r	791	4	33.3	110	2	S50100	transcription-asso
719	4	33.3	102	2	PH1260	Ig heavy chain V r	792	4	33.3	110	2	A55263	hypothetical 12.8k
720	4	33.3	102	2	PH1263	Ig heavy chain V r	793	4	33.3	110	2	TV0102	hypothetical prote
721	4	33.3	102	2	PH1265	Ig heavy chain V r	794	4	33.3	110	2	T30125	probable PE protel
722	4	33.3	102	2	PH1266	Ig heavy chain V r	795	4	33.3	110	2	D70610	hypothetical prote
723	4	33.3	102	2	PH1267	Ig heavy chain V r	796	4	33.3	110	2	T27823	hypothetical prote
724	4	33.3	102	2	PH1275	Ig heavy chain V r	797	4	33.3	110	2	T46071	hypothetical prote
725	4	33.3	102	2	PH1278	Ig heavy chain V r	798	4	33.3	110	2	T48382	hypothetical prote
726	4	33.3	102	2	PH1279	Ig heavy chain V r	799	4	33.3	110	2	G72696	hypothetical prote
727	4	33.3	102	2	PH1281	Ig heavy chain V r	800	4	33.3	111	2	S21925	Ig heavy chain V r
728	4	33.3	102	2	PH1282	Ig heavy chain V r	801	4	33.3	111	2	T16302	hypothetical prote
729	4	33.3	102	2	PH1271	Ig heavy chain V r	802	4	33.3	111	2	A81673	conserved hypotet
730	4	33.3	102	2	PH1272	Ig heavy chain V r	803	4	33.3	111	2	S17200	protein kinase (EC
731	4	33.3	102	2	PH1273	Ig heavy chain V r	804	4	33.3	111	2	D72646	hypothetical prote
732	4	33.3	102	2	H69841	conserved hypotet	805	4	33.3	112	2	E82540	conserved hypotet
733	4	33.3	102	2	S28361	hypothetical prote	806	4	33.3	112	2	S30269	protein hdb precu
734	4	33.3	102	2	F84709	probable glutaredo	807	4	33.3	112	2	F86023	hypothetical prote
735	4	33.3	102	2	T39407	hypothetical prote	808	4	33.3	112	2	S77344	carbon dioxido con
736	4	33.3	103	2	B26167	Ig lambda chain C	809	4	33.3	112	2	A71515	hypothetical prote
737	4	33.3	103	2	B60608	myosin heavy chain	810	4	33.3	112	2	T01515	galactitol utilisz
738	4	33.3	103	2	T65238	parvalbumin - mous	811	4	33.3	112	2	A64976	hypothetical prote
739	4	33.3	103	2	TS1248	ACBP/DBI - duck	812	4	33.3	112	2	C83938	hypothetical prote
740	4	33.3	103	2	G75513	conserved hypotet	813	4	33.3	112	2	C86415	hypothetical prote
741	4	33.3	104	1	GMBO	gastrin precursor	814	4	33.3	113	1	RKWTSS	ribulose-bisphosph
742	4	33.3	104	1	PH1665	Ig heavy chain V r	815	4	33.3	113	1	R6DOP1	acidic ribosomal p
743	4	33.3	104	2	S69899	Ig heavy chain V r	816	4	33.3	113	2	PH1428	Ig heavy chain V r
744	4	33.3	104	2	J00863	hypothetical 11.6k	817	4	33.3	113	2	PH1663	Ig heavy chain V r
745	4	33.3	104	2	A83871	hypothetical prote	818	4	33.3	113	2	PC1281	NS5 protein - hepa
746	4	33.3	104	2	B81787	probable inner mem	819	4	33.3	113	2	S11640	hypothetical prote
747	4	33.3	105	1	L2HU	Ig lambda chain C	820	4	33.3	113	2	A56581	major basic nuclea
748	4	33.3	105	2	S47430	protein kinase (EC	821	4	33.3	113	2	A83429	hypothetical prote
749	4	33.3	105	2	T36213	spdd protein - Str	822	4	33.3	113	2	D83320	hypothetical prote
750	4	33.3	105	2	T49655	hypothetical prote	823	4	33.3	113	2	A83697	hypothetical prote
751	4	33.3	105	2	S27493	modc protein - Bra	824	4	33.3	113	2	T08623	probable sensor xl
752	4	33.3	105	2	S44838	KOZD10.3 protein -	825	4	33.3	113	2	S66805	hypothetical prote
753	4	33.3	106	1	D82734	outer membrane pro	826	4	33.3	113	2	T49533	hypothetical prote
754	4	33.3	106	1	CCDV3S	cytochrome c3 - De	827	4	33.3	113	2	PC2231	hobo-like transpos
755	4	33.3	106	1	K4RBS	Ig kappa-2 chain C	828	4	33.3	114	2	B71088	hypothetical prote
756	4	33.3	106	1	B48354	nonstructural prot	829	4	33.3	114	2	S07898	endothelin 3 - rab
757	4	33.3	106	2	S20774	Ig heavy chain V r	830	4	33.3	114	2	PH1667	Ig heavy chain V r
758	4	33.3	106	2	G20907	Ig kappa-B4 chain	831	4	33.3	114	2	A34567	beta-microseminopr
759	4	33.3	106	2	G72632	hypothetical prote	832	4	33.3	115	1	JN0318	guanylin precursor

541	4	33.3	65	2	S35024	hypothetical prote	614	4	33.3	86	2	F85748	unknown protein en
542	4	33.3	66	2	S65971	yyd protein - Bac	615	4	33.3	86	2	A70005	conserved hypotet
543	4	33.3	66	2	PNO644	hypothetical prote	616	4	33.3	87	1	PF61CP	antifreeze protein
544	4	33.3	67	2	I47395	histone H1 I-1 (cl	617	4	33.3	87	2	B31075	antifreeze protein
545	4	33.3	67	2	G72372	hypothetical prote	618	4	33.3	87	2	C31075	antifreeze protein
546	4	33.3	67	2	A64321	archaeal histone -	619	4	33.3	87	2	T27141	hypothetical prote
547	4	33.3	67	2	A64457	archaeal histone -	620	4	33.3	87	2	S27040	Vat-1 protein - Pa
548	4	33.3	67	2	D64416	archaeal histone -	621	4	33.3	88	1	Q6BP77	gene 6.7 protein -
549	4	33.3	67	2	D64513	archaeal histone.	622	4	33.3	88	2	S60815	M protein precursor
550	4	33.3	68	1	A55457	lactocin S precurs	623	4	33.3	88	2	P85717	hypothetical prote
551	4	33.3	68	2	JH0100	lrpC protein - Bac	624	4	33.3	88	2	A85743	hypothetical prote
552	4	33.3	68	2	C83767	small acid-soluble	625	4	33.3	88	2	A57399	Bm12 protein - Ba
553	4	33.3	69	2	S28195	tissue kallikrein	626	4	33.3	88	2	A83659	hypothetical prote
554	4	33.3	69	2	PC7068	synaptotagmin II p	627	4	33.3	89	2	B69352	Acylphosphatase (a
555	4	33.3	69	2	A82489	cold shock DNA-bin	628	4	33.3	90	2	S78047	DNM-directed RNA p
556	4	33.3	69	2	H69383	conserved hypotet	629	4	33.3	90	2	B71983	hypothetical prote
557	4	33.3	69	2	C86886	50S ribosomal prot	630	4	33.3	91	2	D86503	hypothetical prote
558	4	33.3	70	2	S74243	6-phosphofructo-2-	631	4	33.3	91	2	B96574	hypothetical prote
559	4	33.3	70	2	A47395	histone H1 I-1 (N-	632	4	33.3	91	2	C72119	hypothetical prote
560	4	33.3	70	2	S56769	RNA-directed RNA p	633	4	33.3	92	2	C82174	ribosomal protein
561	4	33.3	71	2	C82807	30S ribosomal prot	634	4	33.3	92	2	PA4620	similar to late em
562	4	33.3	71	2	T45384	ribosomal protein	635	4	33.3	92	2	T34744	hypothetical prote
563	4	33.3	72	2	S28196	tissue kallikrein	636	4	33.3	92	2	T51187	small zinc finger-
564	4	33.3	72	2	PQ0613	ferritin 5 - compe	637	4	33.3	92	2	JU0239	nuclear matrix pro
565	4	33.3	72	2	PQ0614	ferritin 2 - compe	638	4	33.3	93	2	B85537	unknown [imported]
566	4	33.3	72	2	PQ0615	ferritin 1 - compe	639	4	33.3	93	2	T45593	small zinc finger-
567	4	33.3	72	2	S30980	gene 35 protein -	640	4	33.3	94	2	SA9470	hypothetical prote
568	4	33.3	72	2	S75519	hypothetical prote	641	4	33.3	94	2	B84539	hypothetical prote
569	4	33.3	72	2	E69386	hypothetical prote	642	4	33.3	94	2	G36943	chemoreceptor prot
570	4	33.3	72	2	E69523	hypothetical prote	643	4	33.3	94	2	SA43903	hypothetical prote
571	4	33.3	72	2	T48971	hypothetical prote	644	4	33.3	95	2	S68231	PHG22 protein prec
572	4	33.3	73	2	D69319	conserved hypotet	645	4	33.3	95	2	B32830	homeotic protein H
573	4	33.3	73	2	T44814	hypothetical prote	646	4	33.3	95	2	A56644	inverted repeat co
574	4	33.3	73	2	H69035	conserved hypotet	647	4	33.3	95	2	H81119	hypothetical prote
575	4	33.3	73	2	G64443	hypothetical prote	648	4	33.3	95	2	G81901	hypothetical prote
576	4	33.3	73	2	T20911	hypothetical prote	649	4	33.3	95	2	C69905	hypothetical prote
577	4	33.3	73	2	E82788	hypothetical prote	650	4	33.3	96	2	S10069	regulatory protein
578	4	33.3	74	2	E69028	hypothetical prote	651	4	33.3	96	2	T16179	hypothetical prote
579	4	33.3	75	2	T12080	low molecularweig	652	4	33.3	97	1	USBSBW	small acid-soluble
580	4	33.3	75	2	C85636	hypothetical prote	653	4	33.3	97	2	PH0870	Ig heavy chain V r
581	4	33.3	75	2	A64842	ycyc protein - Bsc	654	4	33.3	97	2	I51262	Ig heavy chain V r
582	4	33.3	76	2	T17673	hypothetical prote	655	4	33.3	97	2	A69061	hypothetical prote
583	4	33.3	76	2	B83402	hypothetical prote	656	4	33.3	98	2	PH1429	Ig heavy chain V r
584	4	33.3	76	2	A75309	hypothetical prote	657	4	33.3	98	2	PH1274	Ig heavy chain V r
585	4	33.3	76	2	I39960	hypothetical prote	658	4	33.3	98	2	PH0877	Ig heavy chain V r
586	4	33.3	77	2	H69328	nifu protein (nifu	659	4	33.3	98	2	A30523	Ig heavy chain V-I
587	4	33.3	77	2	D85940	hypothetical prote	660	4	33.3	98	2	PH0871	Ig heavy chain V r
588	4	33.3	77	2	S00970	krclal protein - pl	661	4	33.3	98	2	S24680	Ig heavy chain V1
589	4	33.3	77	2	E82612	hypothetical prote	662	4	33.3	98	2	S46463	Ig heavy chain V1
590	4	33.3	78	2	S61468	p83/100 protein -	663	4	33.3	98	2	S26910	Ig heavy chain V r
591	4	33.3	78	2	G71720	hypothetical prote	664	4	33.3	98	2	S26914	Ig heavy chain V r
592	4	33.3	78	2	S72745	B1177_F2_71 protei	665	4	33.3	98	2	S26918	Ig heavy chain V r
593	4	33.3	78	2	T16325	hypothetical prote	666	4	33.3	98	2	S26938	Ig heavy chain V r
594	4	33.3	79	2	T30118	hypothetical prote	667	4	33.3	98	2	S26909	Ig heavy chain V r
595	4	33.3	80	1	RHIDS	gonadoliberin I pr	668	4	33.3	98	2	S26919	Ig heavy chain V r
596	4	33.3	80	2	F72306	conserved hypotet	669	4	33.3	98	2	S26913	Ig heavy chain V r
597	4	33.3	80	2	D75366	hypothetical prote	670	4	33.3	98	2	S26912	Ig heavy chain V r
598	4	33.3	81	2	T29888	hypothetical prote	671	4	33.3	98	2	S26911	Ig heavy chain V r
599	4	33.3	82	2	S37173	ribosomal protein	672	4	33.3	98	2	S26920	Ig heavy chain V r
600	4	33.3	82	2	G82366	tata protein VC008	673	4	33.3	98	2	S26921	Ig heavy chain V r
601	4	33.3	82	2	T29889	hypothetical prote	674	4	33.3	98	2	S26915	Ig heavy chain V r
602	4	33.3	82	2	T29893	hypothetical prote	675	4	33.3	98	2	S26907	Ig heavy chain V r
603	4	33.3	82	2	T29891	hypothetical prote	676	4	33.3	98	2	I49562	Ig heavy chain V r
604	4	33.3	82	2	T33088	hypothetical prote	677	4	33.3	98	2	S22592	hypothetical prote
605	4	33.3	83	2	C47188	MHC class II histo	678	4	33.3	98	2	T24286	hypothetical prote
606	4	33.3	83	2	T08238	gas-vesicle operon	679	4	33.3	99	1	QOQYBW	ribosomal protein
607	4	33.3	83	2	J01124	serum response fac	680	4	33.3	99	1	H69000	translation initia
608	4	33.3	84	2	A53012	proline rich prote	681	4	33.3	99	2	C29826	hypothetical t1k p
609	4	33.3	84	2	T09540	p83/100 protein -	682	4	33.3	100	2	T08763	hypothetical prote
610	4	33.3	84	2	S61465	p83/100 protein -	683	4	33.3	100	2	B72540	hypothetical prote
611	4	33.3	84	2	S61466	hypothetical prote	684	4	33.3	100	2	B49598	hypothetical prote
612	4	33.3	84	2	T47420	hypothetical prote	685	4	33.3	100	2	S77878	hypothetical prote
613	4	33.3	84	2	E82709	hypothetical prote	686	4	33.3	100	2		



395	5	41.7	1070	2	T25836	hypothetical prote	468	5	41.7	2292	2	S55401	capsid polypepten
396	5	41.7	1070	2	T34385	hypothetical prote	469	5	41.7	2325	2	T02235	acetyl-CoA carboxy
397	5	41.7	1071	2	T04926	starch synthase ho	470	5	41.7	2447	2	T16870	hypothetical prote
398	5	41.7	1073	2	T01955	hypothetical prote	471	5	41.7	2476	2	T34022	zonadhesin - pig
399	5	41.7	1092	2	T33717	carbamoyl-phosphat	472	5	41.7	2541	2	T29340	hypothetical prote
400	5	41.7	1093	2	T50652	AP-3 complex beta-3	473	5	41.7	2796	2	JC4743	fatty-acid synthase
401	5	41.7	1094	2	T50651	AP-3 complex beta-3	474	5	41.7	3131	2	T39553	vacuolar protein s
402	5	41.7	1098	2	S38100	hypothetical prote	475	5	41.7	3131	2	S39842	enantiin synthetas
403	5	41.7	1105	2	T18295	Ap-3 adaptor compl	476	5	41.7	3133	2	S52093	hemocytin - silkw
404	5	41.7	1108	2	T38673	probable transcrip	477	5	41.7	4162	2	T42633	connectin/titin -
405	5	41.7	1113	2	H84105	hypothetical prote	478	5	41.7	4957	2	T03455	ALR protein - huma
406	5	41.7	1119	2	T50995	related to cytoske	479	5	41.7	5262	2	T03454	ALR protein - huma
407	5	41.7	1123	2	S36846	myosin-binding pro	480	5	41.7	6642	2	T29757	protein UNC-89 - C
408	5	41.7	1127	1	E71156	endopeptidase Ia h	481	5	41.7	6839	2	S57242	twitichin (similar
409	5	41.7	1132	2	B82538	ribonuclease E xpr	482	5	41.7	7160	2	T27935	hypothetical prote
410	5	41.7	1138	2	S24614	myosin-binding pro	483	4	33.3	12	2	A33099	163K exoantigen -
411	5	41.7	1150	1	A55289	kinesin-like prote	484	4	33.3	21	1	A35225	conanotkin T - con
412	5	41.7	1171	1	Q0K8FP	pyruvate (flavodox	485	4	33.3	22	2	C39800	seed protein ws-6
413	5	41.7	1172	2	F84572	probable cadmium-t	486	4	33.3	23	2	F61491	potassium channel
414	5	41.7	1196	2	T23832	protein-tyrosine k	487	4	33.3	23	2	PS0446	flavodoxin isoform
415	5	41.7	1199	2	S20969	AF-4 protein, spli	488	4	33.3	30	2	B60291	30K serine protein
416	5	41.7	1210	2	I39410	serine/proline-ric	489	4	33.3	32	2	S21245	H+-transporting At
417	5	41.7	1213	2	A58198	hypothetical prote	490	4	33.3	32	2	S12897	alkaline phosphata
418	5	41.7	1217	2	T25894	web1 protein homol	491	4	33.3	33	2	I22565	R-phycocerythrin ga
419	5	41.7	1224	2	T40765	probable membrane	492	4	33.3	36	2	S57728	phosphatase - Escheric
420	5	41.7	1228	2	S56681	hypothetical prote	493	4	33.3	38	2	S40096	chlorophyll a/b-D1
421	5	41.7	1233	2	S56371	hypothetical prote	494	4	33.3	38	2	T31479	hypothetical prote
422	5	41.7	1240	2	T04193	delta endotoxin -	495	4	33.3	40	2	B69408	hypothetical prote
423	5	41.7	1245	2	T18211	neutonal cell cycl	496	4	33.3	41	2	B54357	muconate cyclisom
424	5	41.7	1251	2	A56677	phospholipase C (E	497	4	33.3	41	2	I68622	hypothetical myell
425	5	41.7	1252	2	S02004	structural polypro	498	4	33.3	42	2	A47116	trifolliotoxin precu
426	5	41.7	1253	1	VHW	1-phosphatidylinos	499	4	33.3	44	1	MMWZK3	K4 protein - vacci
427	5	41.7	1265	2	A34163	insulin receptor-t	500	4	33.3	44	2	S56313	Gut8-2a protein
428	5	41.7	1268	2	B36502	C-terminal domain-	501	4	33.3	45	2	T35464	hypothetical prote
429	5	41.7	1268	2	T31420	transferrin-like p	502	4	33.3	47	2	E47395	histone H1 II-1 (c
430	5	41.7	1274	2	T10729	probable calmoduli	503	4	33.3	48	1	TZAZR2	toxin RPII - sea a
431	5	41.7	1276	2	T02711	chromosomal protei	504	4	33.3	48	2	JN0355	neurotoxin I - sea
432	5	41.7	1290	2	A55094	insulin receptor-t	505	4	33.3	48	2	T12940	hypothetical prote
433	5	41.7	1300	2	A36502	acetyl-CoA carboxy	506	4	33.3	49	2	S39857	ribosomal protein
434	5	41.7	1306	2	S42859	nuclear pore membr	507	4	33.3	49	2	B56448	collilin-like prote
435	5	41.7	1337	2	A53824	hypothetical prote	508	4	33.3	49	2	B86888	50S ribosomal prot
436	5	41.7	1341	2	T17885	zyg-9 protein - Ca	509	4	33.3	50	2	PC4007	hypothetical prote
437	5	41.7	1415	2	T21244	phospholipase C (E	510	4	33.3	50	2	B60718	phospholipase A2 h
438	5	41.7	1418	2	T37264	lin-15B protein -	511	4	33.3	50	2	F84075	hypothetical prote
439	5	41.7	1440	2	T27942	alpha-2-macroglobu	512	4	33.3	51	2	S47401	outer membrane pro
440	5	41.7	1472	2	A26122	sallivary agglutinu	513	4	33.3	52	2	A57336	ropA protein - Khl
441	5	41.7	1473	2	A35186	ribosome receptor,	514	4	33.3	52	2	A38942	chemoreceptor prot
442	5	41.7	1534	2	A56734	cell surface antiq	515	4	33.3	52	2	A69330	ISORF2-like protei
443	5	41.7	1536	2	A43607	hypothetical prote	516	4	33.3	52	2	T00141	hypothetical prote
444	5	41.7	1582	2	T15308	acetyl-CoA carboxy	517	4	33.3	53	2	S66602	aldehyde oxidase (
445	5	41.7	1625	2	T02921	adhesin PI precurs	518	4	33.3	53	2	F47395	histone H1 II-1 (c
446	5	41.7	1637	1	IUYMAP	adhesin PI, group	519	4	33.3	53	2	B47395	histone H1 II-1 (c
447	5	41.7	1635	2	A41480	gravin - human	520	4	33.3	53	2	C47395	histone H1 II-1 (c
448	5	41.7	1684	2	JW0057	acetyl-CoA carboxy	521	4	33.3	53	2	D47395	histone H1 II-2 (c
449	5	41.7	1685	2	T02750	coracle gene prote	522	4	33.3	53	2	G47395	histone H1 II-2 (c
450	5	41.7	1698	2	T13800	1,4-beta-glucanase	523	4	33.3	53	2	A72614	hypothetical prote
451	5	41.7	1711	2	T31337	receptor DSC-205 -	524	4	33.3	54	2	B36943	chemoreceptor prot
452	5	41.7	1723	2	S58880	major merozoite su	525	4	33.3	55	2	T03337	gene e30 protein -
453	5	41.7	1785	2	A45546	hypothetical diver	526	4	33.3	56	2	I56605	I7 alpha-hydroxyta
454	5	41.7	1794	2	T38459	cag island protein	527	4	33.3	56	2	B64368	hypothetical prote
455	5	41.7	1819	2	A71928	hypothetical prote	528	4	33.3	58	2	I79344	interferon alpha O
456	5	41.7	1846	2	T10670	vitellogenin I pre	529	4	33.3	59	2	JT0289	phosphoprotein Pp3
457	5	41.7	1912	2	T29088	cag pathogenicity	530	4	33.3	60	2	B70306	conserved hypotet
458	5	41.7	1937	2	G64585	myosin heavy chain	531	4	33.3	60	2	S53364	mucin 5AC (clone J
459	5	41.7	1938	2	S46773	callose synthase c	532	4	33.3	61	2	T43672	hypothetical prote
460	5	41.7	1963	2	T49914	asht protein - fru	533	4	33.3	62	1	CCB017	ubiquinol--cytochr
461	5	41.7	2144	2	S71490	protein T6D22.14 l	534	4	33.3	62	1	T44918	hypothetical prote
462	5	41.7	2254	2	D86215	gag, pol and env p	535	4	33.3	62	2	T26847	hypothetical prote
463	5	41.7	2272	2	T18572	genome polypepten	536	4	33.3	63	2	E31075	antifreeze protein
464	5	41.7	2290	1	GNNYE	genome polypepten	537	4	33.3	64	2	D31075	antifreeze protein
465	5	41.7	2292	1	GNNYED	genome polypepten	538	4	33.3	64	2	D31075	hypothetical prote
466	5	41.7	2292	1	GNNYEB	capsid polypepten	539	4	33.3	64	2	H82849	hypothetical prote
467	5	41.7	2292	2	S35961	capsid polypepten	540	4	33.3	65	2	G69542	hypothetical prote

249	5	41.7	518	2	S36472	L2 protein - human
250	5	41.7	519	2	S23796	modifier-3 protein
251	5	41.7	519	2	T45038	hypothetical prote
252	5	41.7	525	1	PIWb5	L1 protein - human
253	5	41.7	528	2	E66795	unknown protein F2
254	5	41.7	530	2	D83291	probable MFS trans
255	5	41.7	532	2	S40963	hypothetical prote
256	5	41.7	534	2	C69146	hypothetical prote
257	5	41.7	537	2	D86520	CTP synthetase [im
258	5	41.7	537	2	A72103	CTP synthetase - C
259	5	41.7	537	2	E81568	CTP synthase CP052
260	5	41.7	537	2	E96606	hypothetical prote
261	5	41.7	539	2	I49065	lymphoid-restricti
262	5	41.7	539	2	T02508	hypothetical prote
263	5	41.7	542	2	T29707	hypothetical prote
264	5	41.7	543	2	T00513	cytochrome P450 ho
265	5	41.7	546	2	S36490	L1 protein - human
266	5	41.7	547	2	B56573	nuclear pore compl
267	5	41.7	550	2	C86704	conserved hypotnet
268	5	41.7	553	2	C75318	hypothetical prote
269	5	41.7	556	2	D86262	hypothetical prote
270	5	41.7	556	2	T47511	probable transport
271	5	41.7	557	2	A70480	cardamoyl-phosphat
272	5	41.7	557	2	B83962	hypothetical prote
273	5	41.7	557	2	A00629	protein-tyrosine k
274	5	41.7	558	2	E86016	probable enzyme z4
275	5	41.7	558	2	S75104	hypothetical prote
276	5	41.7	562	2	G84221	NADH dehydrogenase
277	5	41.7	566	2	C70482	repair excision nu
278	5	41.7	567	2	B86481	hypothetical prote
279	5	41.7	570	2	F70332	proline-CRMA synth
280	5	41.7	572	1	FOHYH	retrovirus-related
281	5	41.7	573	2	T21653	hypothetical prote
282	5	41.7	573	2	S66710	probable membrane
283	5	41.7	574	1	S76132	protein kinase pkn
284	5	41.7	578	2	T33916	hypothetical prote
285	5	41.7	580	2	C96597	Rubisco subunit b1
286	5	41.7	581	2	E75383	conserved hypotnet
287	5	41.7	583	2	S01301	lamin I-III - Afri
288	5	41.7	583	2	T30131	hypothetical prote
289	5	41.7	585	2	E85809	hypothetical prote
290	5	41.7	587	2	F85084	probable ath1a-11
291	5	41.7	599	2	D81367	probable pyruvate
292	5	41.7	601	1	A64222	heat shock protein
293	5	41.7	603	2	S06059	gene ND1 intron 4
294	5	41.7	604	2	S54032	probable amino aci
295	5	41.7	604	2	S25203	stmr protein - Str
296	5	41.7	605	2	E69153	cadmium efflux ATP
297	5	41.7	609	2	S45930	probable amino aci
298	5	41.7	612	2	G71972	threonine--tRNA L1
299	5	41.7	612	2	C64535	threonine--tRNA L1
300	5	41.7	615	2	T37981	probable exocyst c
301	5	41.7	622	2	A64980	hypothetical 68.5K
302	5	41.7	634	1	MIWL51	E1 protein - human
303	5	41.7	634	2	I49642	estrogen-responsiv
304	5	41.7	637	2	S74677	hypothetical prote
305	5	41.7	641	2	T03095	homeoprotein Sail
306	5	41.7	643	1	DIEC36	DNA-directed DNA p
307	5	41.7	643	1	G85344	hypothetical prote
308	5	41.7	646	2	G69871	hypothetical prote
309	5	41.7	649	2	JN0809	drebrin E (clone 9
310	5	41.7	651	2	T15624	hypothetical prote
311	5	41.7	651	2	T21175	hypothetical prote
312	5	41.7	651	2	T14763	hypothetical prote
313	5	41.7	654	1	HMIVC8	hemagglutinin prec
314	5	41.7	655	2	S07412	hemagglutinin-este
315	5	41.7	662	2	T20570	hypothetical prote
316	5	41.7	668	2	C71365	probable excludnle
317	5	41.7	668	2	H86729	exinuclease ABC s
318	5	41.7	670	2	T34548	hypothetical prote
319	5	41.7	671	2	A38109	autolysin - Entero
320	5	41.7	682	2	C84295	UDP-sugar hydrolas
321	5	41.7	683	2	E86358	hypothetical prote
322	5	41.7	691	2	H81313	ATP-dependent DNA
323	5	41.7	700	2	T13577	hypothetical prote
324	5	41.7	701	1	S33709	DHR39-short protei
325	5	41.7	707	2	JC2218	procollagen C-endo
326	5	41.7	711	2	A85352	cadmalin-transport
327	5	41.7	713	2	A28706	calpastatin, cardl
328	5	41.7	730	1	BMHU1	procollagen C-endo
329	5	41.7	739	2	T29407	hypothetical prote
330	5	41.7	746	2	T19287	hypothetical prote
331	5	41.7	748	2	T49633	glucan 1,4-alpha-g
332	5	41.7	748	2	S4505	hypothetical prote
333	5	41.7	750	2	H70628	probable pknng prot
334	5	41.7	752	1	S51866	HPI1 protein - yea
335	5	41.7	755	2	S42462	structural polypro
336	5	41.7	763	2	T44735	serine/threonine-s
337	5	41.7	767	2	S63220	probable membrane
338	5	41.7	768	2	T37601	probable transcript
339	5	41.7	771	2	T26176	hypothetical prote
340	5	41.7	775	2	E83400	pyroloquinoline q
341	5	41.7	781	1	S64576	hypothetical prote
342	5	41.7	792	2	S20554	pyruvate, water dik
343	5	41.7	792	2	E85778	phosphoenolpyruvat
344	5	41.7	794	2	T37989	DNA mismatch repair
345	5	41.7	806	2	H70647	probable NADH dehy
346	5	41.7	808	1	S33708	nuclear steroid/ch
347	5	41.7	811	2	PN0689	connectin I - chic
348	5	41.7	815	2	H96494	protein F7F22.2 [i
349	5	41.7	816	2	T17257	hypothetical prote
350	5	41.7	817	2	S51342	verpitolin - yeast
351	5	41.7	822	2	G69010	conserved hypotnet
352	5	41.7	823	1	A58788	procollagen C-endo
353	5	41.7	829	2	T33283	hypothetical prote
354	5	41.7	834	2	E69380	methyl-accepting c
355	5	41.7	835	2	S11442	livaslin Yersinia
356	5	41.7	835	2	C86444	hypothetical prote
357	5	41.7	842	2	S49124	carbon-monoxide de
358	5	41.7	865	2	A47282	calcium-binding pr
359	5	41.7	871	2	H81430	translaton initia
360	5	41.7	871	2	D86355	protein T16E15.12
361	5	41.7	873	2	A47283	calphostin - fruit
362	5	41.7	877	2	E81847	pyruvate dehydroge
363	5	41.7	887	2	E81094	S501 protein - yea
364	5	41.7	899	2	S51341	genome polypotein
365	5	41.7	900	1	GNNYV	hypothetical prote
366	5	41.7	912	2	H71931	YopC/gen secretion
367	5	41.7	919	2	E65578	type III secretion
368	5	41.7	919	2	F72045	cell division prot
369	5	41.7	920	2	T43263	probable excludnle
370	5	41.7	925	2	T33732	hypothetical prote
371	5	41.7	929	2	G71420	hypothetical prote
372	5	41.7	930	2	A25923	progesterone recep
373	5	41.7	931	2	D86222	protein F7G19.9 [i
374	5	41.7	933	1	QRHUP	progesterone recep
375	5	41.7	940	2	T01834	hypothetical prote
376	5	41.7	951	2	T08987	probable cadmium-t
377	5	41.7	954	2	S57108	hypothetical prote
378	5	41.7	958	2	S47179	hypothetical prote
379	5	41.7	969	2	A75634	McB-related prote
380	5	41.7	980	2	S71090	peroxisome biogene
381	5	41.7	982	2	T43699	DNA mismatch repair
382	5	41.7	986	1	B58788	procollagen C-endo
383	5	41.7	990	2	I51618	nucleolar phosphop
384	5	41.7	990	2	T12678	hypothetical prote
385	5	41.7	991	2	I49540	procollagen C-endo
386	5	41.7	998	2	H75005	ATP-dependent prot
387	5	41.7	1000	2	B72385	hypothetical prote
388	5	41.7	1001	2	T00532	myosin-B - Toxopla
389	5	41.7	1001	2	T29105	FUN12 protein - ye
390	5	41.7	1002	2	S70292	probable mbtd prot
391	5	41.7	1004	2	H70673	cytodiference acces
392	5	41.7	1018	1	S73720	leukocyte surface
393	5	41.7	1021	2	I39207	hypothetical prote
394	5	41.7	1025	2	E86355	

103	5	41.7	276	2	T29894	hypothetical prote
104	5	41.7	287	2	E75159	hypothetical prote
105	5	41.7	288	2	A81009	hypothetical prote
106	5	41.7	289	2	C70400	ferredoxin oxidore
107	5	41.7	291	2	S42420	replication protei
108	5	41.7	293	2	T09758	myb-related protei
109	5	41.7	293	2	A82690	phosphatidyserine
110	5	41.7	296	2	S46018	probable membrane
111	5	41.7	296	2	G84747	Ar-hook DNA-bindin
112	5	41.7	298	2	T36900	probable integral
113	5	41.7	299	2	A82063	uroporphyrin-III C
114	5	41.7	302	2	A96841	hypothetical prote
115	5	41.7	305	2	JN0518	acylttransferase (E
116	5	41.7	305	2	F83966	dehydrocorate deh
117	5	41.7	307	1	S75769	hypothetical prote
118	5	41.7	308	2	E75409	probable antilbroti
119	5	41.7	309	1	A25776	urate oxidase (EC
120	5	41.7	311	2	H84746	hypothetical prote
121	5	41.7	314	2	S22513	ethylene-forming e
122	5	41.7	315	2	T25171	hypothetical prote
123	5	41.7	315	2	T26422	hypothetical prote
124	5	41.7	318	2	T00112	glycosyltransferas
125	5	41.7	321	2	U80772	37k glycoprotein -
126	5	41.7	321	2	D70070	transcription regu
127	5	41.7	323	2	A31351	probable transmemb
128	5	41.7	325	2	D70666	probable modd prot
129	5	41.7	327	2	JC4195	36k hydrophobic nu
130	5	41.7	327	2	H82554	conserved hypothet
131	5	41.7	335	2	T31559	hypothetical prote
132	5	41.7	335	2	T31560	hypothetical prote
133	5	41.7	335	2	T31561	hypothetical prote
134	5	41.7	335	2	G75309	lyta protein - Del
135	5	41.7	335	2	H75518	probable cytochrom
136	5	41.7	335	2	T37262	probable tyrosine
137	5	41.7	336	1	B38274	Y box-binding prot
138	5	41.7	337	2	T46742	ornithine carboxamoy
139	5	41.7	339	2	G75529	peptide ABC transp
140	5	41.7	340	2	D86837	ornithine carboxamoy
141	5	41.7	342	2	T45456	probable membrane
142	5	41.7	343	2	T14245	NADH dehydrogenase
143	5	41.7	343	2	H82171	probable multidrug
144	5	41.7	346	2	S56665	tryptophan synthas
145	5	41.7	355	2	C70457	hypothetical prote
146	5	41.7	356	1	XNECHC	histidinol-phospha
147	5	41.7	356	2	D85827	histidinol-phospha
148	5	41.7	359	1	XNEBHC	histidinol-phospha
149	5	41.7	360	2	T26037	hypothetical prote
150	5	41.7	360	2	S32695	Wnt-2 protein - Ca
151	5	41.7	362	2	S24551	protein-tyrosine k
152	5	41.7	366	1	MNXRSD	sigma NS protein -
153	5	41.7	366	1	MNXRSD	sigma NS protein -
154	5	41.7	366	1	MNXRST	sigma NS protein -
155	5	41.7	366	2	C27401	sigma NS protein -
156	5	41.7	366	2	S53073	hypothetical prote
157	5	41.7	367	1	A46355	site-specific DNA-
158	5	41.7	367	2	T18185	probable site-spec
159	5	41.7	367	2	S46336	enu polypeptide -
160	5	41.7	368	2	S06058	NADH dehydrogenase
161	5	41.7	375	2	T51333	transcription fact
162	5	41.7	376	2	S04497	surface antigen PA
163	5	41.7	379	2	A45443	ubulin--tyrosine
164	5	41.7	380	2	D84214	NADH dehydrogenase
165	5	41.7	386	2	S64614	probable membrane
166	5	41.7	389	2	D84068	RNA helicase BH334
167	5	41.7	390	2	F65048	multidrug resistanc
168	5	41.7	390	2	F85916	multidrug resistanc
169	5	41.7	391	2	I51097	thyroid hormone re
170	5	41.7	394	2	T18752	hypothetical prote
171	5	41.7	395	2	D83782	aspartate aminotra
172	5	41.7	402	2	JH0403	procollagen I C-pr
173	5	41.7	410	2	D70599	hypothetical prote
174	5	41.7	415	2	T20335	hypothetical prote
175	5	41.7	416	2	G69748	conserved hypothet
176	5	41.7	418	2	F83986	transporter BH2694
177	5	41.7	421	2	D81098	gcpe protein NM813
178	5	41.7	421	2	H81843	conserved hypothet
179	5	41.7	421	2	T14773	hypothetical prote
180	5	41.7	422	2	G70984	probable papas pro
181	5	41.7	423	2	H84257	Htr13 transducer I
182	5	41.7	423	2	T44258	transducer protein
183	5	41.7	423	2	S73020	hypothetical prote
184	5	41.7	428	2	S28468	mannose-1-phosphat
185	5	41.7	432	1	A41056	brachyury homolog
186	5	41.7	435	2	T19687	hypothetical prote
187	5	41.7	440	2	D64616	N-acetylmuramoyl-L
188	5	41.7	442	2	D71481	probable trigger f
189	5	41.7	443	2	E71135	hypothetical prote
190	5	41.7	445	2	D71977	phosphoglucosamine
191	5	41.7	445	2	C38537	urec protein - Hel
192	5	41.7	451	2	B56558	retinoic acid rece
193	5	41.7	451	2	B96774	hypothetical prote
194	5	41.7	465	2	T38962	hypothetical prote
195	5	41.7	453	2	S40943	hypothetical prote
196	5	41.7	459	2	A41977	retinoic acid rece
197	5	41.7	461	2	T42513	1D-myo-inositol-tr
198	5	41.7	462	2	B86262	hypothetical prote
199	5	41.7	464	2	A56558	retinoic acid rece
200	5	41.7	465	2	H82345	mannose-1-phosphat
201	5	41.7	469	2	H71897	probable probable
202	5	41.7	470	2	T15946	hypothetical prote
203	5	41.7	471	2	JC1403	glutamate--ammonia
204	5	41.7	473	2	D70405	glutamate--ammonia
205	5	41.7	473	2	S75141	glutamate--ammonia
206	5	41.7	474	2	H69761	conserved hypothet
207	5	41.7	474	2	T10271	capsid-associated
208	5	41.7	475	2	D44997	conserved hypothet
209	5	41.7	475	2	D86209	protein F22G5.18 l
210	5	41.7	475	2	A85923	hypothetical prote
211	5	41.7	476	2	B44997	merozoite surface
212	5	41.7	476	2	T25736	hypothetical prote
213	5	41.7	477	2	A75545	amidiophosphoribosy
214	5	41.7	478	2	A32555	major merozoite su
215	5	41.7	482	2	A44997	merozoite surface
216	5	41.7	482	2	T32564	hypothetical prote
217	5	41.7	484	2	D65230	hypothetical 52.9
218	5	41.7	484	2	A86116	hypothetical prote
219	5	41.7	488	1	QOEBHS	alkaline exonuclea
220	5	41.7	488	2	T44030	alkaline exonuclea
221	5	41.7	488	2	T44215	alkaline exonuclea
222	5	41.7	488	2	F86209	protein F22G5.14 l
223	5	41.7	490	2	P00164	sucrose phosphoryl
224	5	41.7	492	2	T43859	uroporphyrinogen I
225	5	41.7	493	2	T12044	probable legumain
226	5	41.7	496	2	T30976	hypothetical prote
227	5	41.7	498	2	D64878	probable glutamate
228	5	41.7	498	2	G02421	nicotinic acetylch
229	5	41.7	498	2	G85755	glutamate--ammonia
230	5	41.7	500	2	S77243	hypothetical prote
231	5	41.7	501	2	T27513	hypothetical prote
232	5	41.7	504	2	T34995	probable integral
233	5	41.7	504	2	G02474	interferon regulat
234	5	41.7	505	2	T01667	gag polyprotein
235	5	41.7	505	2	B64560	poly E-rich protei
236	5	41.7	505	2	B46629	mucin 6, gastric (
237	5	41.7	507	2	G01614	zinc finger protei
238	5	41.7	508	2	T19350	hypothetical prote
239	5	41.7	509	1	S04346	steroid 17alpha-mo
240	5	41.7	509	2	S36572	LI protein - human
241	5	41.7	510	2	S53970	amidophosphoribosy
242	5	41.7	512	2	T33463	probable serine ca
243	5	41.7	514	1	PIWR8	LI protein - human
244	5	41.7	514	2	PIWR47	LI protein - human
245	5	41.7	516	1	PIWR5	LI protein - human
246	5	41.7	516	2	T09364	cytochrome p450 ho
247	5	41.7	517	2	S36496	LI protein - human
248	5	41.7	517	2	S36543	LI protein - human

GenCore version 4.5.  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:34:40 ; Search time 33.87 Seconds  
(without alignments)  
26.988 Million cell updates/sec

Title: US-09-372-036-30  
Perfect score: 12  
Sequence: 1 SRPVAPTOEVKK 12

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	100.0	484	2 A1487	protein p60 precu
2	7	58.3	227	2 D81259	probable tomb tran
3	6	50.0	102	2 E83231	hypothetical prote
4	6	50.0	116	2 T37186	hypothetical prote
5	6	50.0	155	2 S35162	SRF-21 protein - p
6	6	50.0	187	2 JC4806	core protein g - p
7	6	50.0	218	2 C68847	transcription regu
8	6	50.0	288	2 F71504	hypothetical prote
9	6	50.0	293	2 F75466	hypothetical prote
10	6	50.0	447	2 T34992	probable lipoprote
11	6	50.0	453	1 F64623	amidase - Helicoba
12	6	50.0	453	2 A71891	glu-tRNA amidotran
13	6	50.0	456	2 H85574	hypothetical prote
14	6	50.0	460	2 T33087	hypothetical prote
15	6	50.0	617	2 S42719	actin-binding prote
16	6	50.0	637	2 S66953	hypothetical prote
17	6	50.0	748	2 H84913	probable ATP-depen
18	6	50.0	985	1 VCLJSP	env polypeptide -
19	6	50.0	1133	2 T30302	P-type ATPase - Te
20	6	50.0	1230	2 E64664	outer membrane pro
21	5	41.7	75	2 E75264	hypothetical prote
22	5	41.7	82	2 T29802	hypothetical prote
23	5	41.7	88	2 T34429	hypothetical prote
24	5	41.7	102	2 B56581	major basic nuclea
25	5	41.7	103	1 K4RB	Ig kappa-B4 chain
26	5	41.7	104	2 F53275	Ig kappa-1 chain C
27	5	41.7	112	2 T14788	hypothetical prote
28	5	41.7	114	2 T08411	hypothetical prote
29	5	41.7	115	2 B49094	methylnalonyl-CoA

30	5	41.7	117	2 A72293	conserved hypothet
31	5	41.7	124	2 B66557	Dnak suppressor (I
32	5	41.7	124	2 A71518	probable dnak supp
33	5	41.7	124	2 G72066	dnak suppressor
34	5	41.7	124	2 T24876	hypothetical prote
35	5	41.7	126	2 I54454	lymphocyte antigen
36	5	41.7	131	2 D86684	prophage p11 prote
37	5	41.7	132	2 B81600	conserved hypothet
38	5	41.7	133	2 T50083	probable cyclidine
39	5	41.7	134	2 B60497	MHC class II histo
40	5	41.7	134	2 A81676	conserved hypothet
41	5	41.7	141	2 E72580	hypothetical prote
42	5	41.7	147	2 G72384	conserved hypothet
43	5	41.7	147	2 T33327	hypothetical prote
44	5	41.7	157	2 T25638	hypothetical prote
45	5	41.7	161	2 C96568	hypothetical prote
46	5	41.7	165	2 S61230	cytochrome-c biosy
47	5	41.7	170	2 S11770	lma protein - Lis
48	5	41.7	174	2 S07291	repressor protein
49	5	41.7	175	2 D86639	hypothetical prote
50	5	41.7	175	2 T50849	hypothetical prote
51	5	41.7	177	1 S57789	hypothetical prote
52	5	41.7	178	2 S56300	hypothetical prote
53	5	41.7	179	2 G84789	hypothetical prote
54	5	41.7	185	2 H86513	elongation factor
55	5	41.7	185	2 G72110	translational elonga
56	5	41.7	185	2 D96572	protein F12M16.6 [
57	5	41.7	190	2 H75162	anthranilate synth
58	5	41.7	193	2 H66404	probable lipid tra
59	5	41.7	196	2 A41050	repressor protein
60	5	41.7	199	2 F86832	hypothetical prote
61	5	41.7	200	2 F64217	ribosomal protein
62	5	41.7	201	1 IMBPSB	immunoty protein -
63	5	41.7	201	2 S69748	hypothetical prote
64	5	41.7	206	2 A71326	probable V-type AT
65	5	41.7	206	2 H83587	conserved hypothet
66	5	41.7	209	2 S04827	gene 40A protein -
67	5	41.7	211	2 T22122	hypothetical prote
68	5	41.7	217	2 S10212	late 33k protein -
69	5	41.7	217	2 T37859	probable transcrip
70	5	41.7	218	2 G71260	probable ribulose-
71	5	41.7	221	2 H86860	protein serine/thr
72	5	41.7	223	2 G82478	hypothetical prote
73	5	41.7	224	2 A31459	MHC class II histo
74	5	41.7	227	2 T12797	immunoty protein d
75	5	41.7	228	2 H71214	hypothetical prote
76	5	41.7	228	2 F75390	hypothetical prote
77	5	41.7	239	2 A27207	lysine kallikrein
78	5	41.7	241	2 S75154	urease accessory p
79	5	41.7	245	2 F71474	probable muramidase
80	5	41.7	246	2 I40763	hypothetical prote
81	5	41.7	246	2 B81373	hypothetical prote
82	5	41.7	250	2 S62838	hypothetical prote
83	5	41.7	251	2 T45707	CHLOROPHYLL A-B BI
84	5	41.7	251	2 C84036	succinate dehydrog
85	5	41.7	252	2 A86449	hypothetical prote
86	5	41.7	253	2 E71198	hypothetical prote
87	5	41.7	253	2 A75159	hypothetical prote
88	5	41.7	254	2 T33323	protein-1-isospar
89	5	41.7	260	2 T18909	hypothetical prote
90	5	41.7	261	1 HLHDB	MHC class II histo
91	5	41.7	261	1 HLHUC	MHC class II histo
92	5	41.7	261	1 B37044	MHC class II histo
93	5	41.7	261	2 I54480	HLA DQ-beta - huma
94	5	41.7	261	2 T51222	hypothetical prote
95	5	41.7	263	2 C69887	hypothetical prote
96	5	41.7	264	2 C37386	hypothetical prote
97	5	41.7	264	2 S01922	hypothetical prote
98	5	41.7	267	2 T32063	hypothetical prote
99	5	41.7	268	2 D29312	hypothetical prote
100	5	41.7	269	2 I54432	MHC class II histo
101	5	41.7	271	2 T16590	MHC class II histo
102	5	41.7	276	2 S75249	hypothetical prote

conserved hypothet  
Dnak suppressor (I  
probable dnak supp  
dnak suppressor  
hypothetical prote  
lymphocyte antigen  
prophage p11 prote  
conserved hypothet  
probable cyclidine  
MHC class II histo  
conserved hypothet  
hypothetical prote  
conserved hypothet  
hypothetical prote  
hypothetical prote  
cytochrome-c biosy  
lma protein - Lis  
repressor protein  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
elongation factor  
translational elonga  
protein F12M16.6 [  
anthranilate synth  
probable lipid tra  
repressor protein  
hypothetical prote  
ribosomal protein  
immunoty protein -  
hypothetical prote  
probable V-type AT  
conserved hypothet  
gene 40A protein -  
hypothetical prote  
late 33k protein -  
probable transcrip  
probable ribulose-  
protein serine/thr  
hypothetical prote  
MHC class II histo  
immunoty protein d  
hypothetical prote  
hypothetical prote  
lysine kallikrein  
urease accessory p  
probable muramidase  
hypothetical prote  
hypothetical prote  
CHLOROPHYLL A-B BI  
succinate dehydrog  
hypothetical prote  
hypothetical prote  
protein-1-isospar  
hypothetical prote  
hypothetical prote  
MHC class II histo  
MHC class II histo  
MHC class II histo  
HLA DQ-beta - huma  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
MHC class II histo  
MHC class II histo  
hypothetical prote  
esterase s110992 -

**THIS PAGE BLANK (USPTO)**

---

PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150584.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154779.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 28-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159826.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 20-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144684.	PR 14-OCT-1999;	99US-0159637.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161992.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 04-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		

Query Match 41.7%; Score 5; DB 21; Length 158;  
Best Local Similarity 100.0%; Pred. No. 34e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 rpvap 6  
| | | |  
Db 6 lrpvap 10

Search completed: August 15, 2001, 12:35:47  
Job time: 172 sec

PR 10-AUG-1999; 990S-0148171.  
PR 11-AUG-1999; 990S-0148319.  
PR 12-AUG-1999; 990S-0148341.  
PR 13-AUG-1999; 990S-0148565.  
PR 13-AUG-1999; 990S-0148684.  
PR 16-AUG-1999; 990S-0149368.  
PR 17-AUG-1999; 990S-0149175.  
PR 18-AUG-1999; 990S-0149426.  
PR 20-AUG-1999; 990S-0149722.  
PR 20-AUG-1999; 990S-0149723.  
PR 20-AUG-1999; 990S-0149929.  
PR 23-AUG-1999; 990S-0149902.  
PR 23-AUG-1999; 990S-0149930.  
PR 25-AUG-1999; 990S-0150566.  
PR 26-AUG-1999; 990S-0150884.  
PR 27-AUG-1999; 990S-0151065.  
PR 27-AUG-1999; 990S-0151066.  
PR 27-AUG-1999; 990S-0151080.  
PR 30-AUG-1999; 990S-0151303.  
PR 31-AUG-1999; 990S-0151438.  
PR 01-SEP-1999; 990S-0151930.  
PR 07-SEP-1999; 990S-0152363.  
PR 10-SEP-1999; 990S-0153070.  
PR 13-SEP-1999; 990S-0153758.  
PR 15-SEP-1999; 990S-0154018.  
PR 16-SEP-1999; 990S-0154039.  
PR 20-SEP-1999; 990S-0154779.  
PR 22-SEP-1999; 990S-0155139.  
PR 23-SEP-1999; 990S-0155486.  
PR 24-SEP-1999; 990S-0155659.  
PR 28-SEP-1999; 990S-0156458.  
PR 29-SEP-1999; 990S-0156596.  
PR 04-OCT-1999; 990S-0157117.  
PR 05-OCT-1999; 990S-0157753.  
PR 06-OCT-1999; 990S-0157865.  
PR 07-OCT-1999; 990S-0158029.  
PR 08-OCT-1999; 990S-0158232.  
PR 12-OCT-1999; 990S-0158369.  
PR 13-OCT-1999; 990S-0158923.  
PR 13-OCT-1999; 990S-0159294.  
PR 13-OCT-1999; 990S-0159295.  
PR 14-OCT-1999; 990S-0159329.  
PR 14-OCT-1999; 990S-0159330.  
PR 14-OCT-1999; 990S-0159331.  
PR 14-OCT-1999; 990S-0159637.  
PR 14-OCT-1999; 990S-0159638.  
PR 18-OCT-1999; 990S-0159584.  
PR 21-OCT-1999; 990S-0160741.  
PR 21-OCT-1999; 990S-0160767.  
PR 21-OCT-1999; 990S-0160768.  
PR 21-OCT-1999; 990S-0160770.  
PR 21-OCT-1999; 990S-0160814.  
PR 21-OCT-1999; 990S-0160815.  
PR 22-OCT-1999; 990S-0160980.  
PR 22-OCT-1999; 990S-0160981.  
PR 22-OCT-1999; 990S-0160989.  
PR 25-OCT-1999; 990S-0161404.  
PR 25-OCT-1999; 990S-0161405.  
PR 25-OCT-1999; 990S-0161406.  
PR 26-OCT-1999; 990S-0161359.  
PR 26-OCT-1999; 990S-0161360.  
PR 26-OCT-1999; 990S-0161361.  
PR 28-OCT-1999; 990S-0161920.  
PR 28-OCT-1999; 990S-0161993.  
PR 28-OCT-1999; 990S-0161992.  
PR 29-OCT-1999; 990S-0162142.

Query Match 41.7%; Score 5; DB 21; Length 152;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 3 PVAPT 7  
|||||

Db 95 pvapt 99  
RESULT 50  
AAG12578  
ID AAG12578 standard; Protein; 158 AA.  
XX  
AC AAG12578;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Zea mays protein fragment SEQ ID NO: 11746.  
XX  
DE Zea mays protein fragment SEQ ID NO: 11746.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence; corn.  
OS Zea mays subsp. mays.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 990S-0121825.  
PR 05-MAR-1999; 990S-0123180.  
PR 09-MAR-1999; 990S-0123548.  
PR 23-MAR-1999; 990S-0125788.  
PR 25-MAR-1999; 990S-0126264.  
PR 29-MAR-1999; 990S-0126765.  
PR 01-APR-1999; 990S-0127462.  
PR 06-APR-1999; 990S-0128234.  
PR 08-APR-1999; 990S-0128714.  
PR 16-APR-1999; 990S-0129845.  
PR 19-APR-1999; 990S-0130077.  
PR 21-APR-1999; 990S-0130077.  
PR 23-APR-1999; 990S-0130510.  
PR 23-APR-1999; 990S-0130891.  
PR 28-APR-1999; 990S-0131449.  
PR 30-APR-1999; 990S-0132048.  
PR 30-APR-1999; 990S-0132407.  
PR 04-MAY-1999; 990S-0132484.  
PR 05-MAY-1999; 990S-0132485.  
PR 06-MAY-1999; 990S-0132486.  
PR 06-MAY-1999; 990S-0132487.  
PR 07-MAY-1999; 990S-0132863.  
PR 11-MAY-1999; 990S-0134256.  
PR 14-MAY-1999; 990S-0134218.  
PR 14-MAY-1999; 990S-0134219.  
PR 14-MAY-1999; 990S-0134221.  
PR 14-MAY-1999; 990S-0134370.  
PR 18-MAY-1999; 990S-0134768.  
PR 19-MAY-1999; 990S-0134941.  
PR 20-MAY-1999; 990S-0135124.  
PR 21-MAY-1999; 990S-0135353.  
PR 24-MAY-1999; 990S-0135629.  
PR 25-MAY-1999; 990S-0136021.  
PR 25-MAY-1999; 990S-0136392.  
PR 27-MAY-1999; 990S-0136782.  
PR 28-MAY-1999; 990S-0137222.  
PR 01-JUN-1999; 990S-0137528.  
PR 03-JUN-1999; 990S-0137502.  
PR 04-JUN-1999; 990S-0137724.  
PR 07-JUN-1999; 990S-0138094.  
PR 08-JUN-1999; 990S-0138540.  
PR 10-JUN-1999; 990S-0138847.  
PR 14-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139452.  
PR 16-JUN-1999; 990S-0139453.  
PR 17-JUN-1999; 990S-0139492.  
PR 18-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139455.

Query Match 41.7%; Score 5; DB 21; Length 151;  
Best Local Similarity 100.0%; Pred. No. 3,3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 npvap 6  
      ||||  
Db 50 npvap 54

RESULT 49  
AAG54453  
ID AAG54453 standard; Protein; 152 AA.

XX AAG54453;

DT 18-OCT-2000 (first entry)

DE Zea mays protein fragment SEQ ID NO: 69433.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.

XX Zea mays subsp. mays.

PN EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132408.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 24-JUN-1999; 99US-0140354.

PR 28-JUN-1999; 99US-0140695.

PR 29-JUN-1999; 99US-0140823.

PR 30-JUN-1999; 99US-0140991.

PR 01-JUL-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 19-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 22-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 28-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.



```

PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147935.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149722.
PR 20-AUG-1999; 9905-0149723.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151338.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 15-SEP-1999; 9905-0154039.
PR 16-SEP-1999; 9905-0154779.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160815.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 25-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 26-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

```

Query Match 41.78; Score 5; DB 21; Length 148;  
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 APR0E 9  
 |||||  
 Db 93 aptge 97

RESULT 48  
 AAB43699  
 ID AAB43699 standard; Protein; 151 AA.

AC AAB43699;

DT 08-FEB-2001 (first entry)

DE Human cancer associated protein sequence SEQ ID NO:1144.

Human: Cancer associated gene; cancer antigen; detection; cancer;  
 diagnosis; cytostatic; proliferative; vulnerrary; immunomodulator;  
 antidiabetic; antilastmatic; antineumatic; antiarthritic; antiviral;  
 antinflammatory; antithyroid; antiallergic; antibacterial; cardiant;  
 dermatological; neuroprotective; thrombolytic; coagulant; nocrotropic;  
 vasotropic; antiporietic; antiangiogenic; gene therapy; inflammation;  
 immune disorder; haematopoietic cell disorder; autoimmune disorder;  
 allergic reaction; graft versus host disease; organ rejection;  
 haemostatic; thrombolytic; cardiovascular disorder; infection;  
 neurological disease; drug screening.

OS Homo sapiens.

PN WO200055350-A1.

PD 21-SEP-2000.

PF 08-MAR-2000; 2000MO-US05882.

PR 12-MAR-1999; 9905-0124270.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM;

DR WPI; 2000-587533/55.

DR N-PSDB; AAC77908.

PT Novel isolated nucleic acids comprising sequences encoding peptides  
 useful for treating or diagnosing e.g. cancer -

XX Claim 11; Page 1760-1761; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given  
 in AAB43398 to AAB44239. The proteins can have activities based on the  
 tissues and cells the genes are expressed in. Example of activities  
 include: cytostatic; proliferative; vulnerrary; immunomodulator;  
 antidiabetic; antilastmatic; antineumatic; antiarthritic;  
 antinflammatory; antithyroid; antiallergic; antibacterial; antiviral;  
 dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
 nocrotropic; vasotropic; antiporietic and antiangiogenic. The  
 polynucleotides and polypeptides can be used for preventing, treating or  
 ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 the present invention may be used to treat immune disorders by activating  
 or inhibiting the proliferation, differentiation or mobilisation of  
 immune cells, to treat disorders of haematopoietic cells, autoimmune  
 disorders, allergic reactions, graft versus host disease and organ  
 rejection, modulate haemostatic or thrombolytic activity, modulate  
 inflammation, cancers, cardiovascular disorders, neurological disease and  
 bacterial or viral infections. The peptides, nucleotides, antibodies,  
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
 AAC78457 and AAB44240 represent sequences used in the exemplification of  
 the present invention.

XX Sequence 151 AA;

PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 41.7%; Score 5; DB 21; Length 148;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APTOE 9  
Db 93 aptge 97

RESULT 47  
AAG53842

ID AAG53842 standard; Protein; 148 AA.

AC AAG53842;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 68587.

XX protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.

PR 24-MAY-1999; 9905-0135629.  
PR 25-MAY-1999; 9905-0136021.  
PR 27-MAY-1999; 9905-0136392.  
PR 28-MAY-1999; 9905-0136782.  
PR 01-JUN-1999; 9905-0137222.  
PR 03-JUN-1999; 9905-0137528.  
PR 04-JUN-1999; 9905-0137502.  
PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139452.  
PR 16-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139492.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.  
PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140695.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141842.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142977.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0143624.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144652.  
PR 20-JUL-1999; 9905-0144884.  
PR 20-JUL-1999; 9905-0144884.  
PR 21-JUL-1999; 9905-0144814.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145085.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145218.  
PR 23-JUL-1999; 9905-0145224.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 27-JUL-1999; 9905-0145919.  
PR 28-JUL-1999; 9905-0145951.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
  
PR 02-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0147038.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147392.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.  
PR 06-AUG-1999; 9905-0147416.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147935.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148565.  
PR 13-AUG-1999; 9905-0148684.  
PR 16-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149375.  
PR 18-AUG-1999; 9905-0149426.  
PR 20-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149723.  
PR 20-AUG-1999; 9905-0149929.  
PR 23-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150566.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151065.  
PR 27-AUG-1999; 9905-0151066.  
PR 30-AUG-1999; 9905-0151080.  
PR 30-AUG-1999; 9905-0151303.  
PR 31-AUG-1999; 9905-015138.  
PR 01-SEP-1999; 9905-0151930.  
PR 07-SEP-1999; 9905-0152363.  
PR 10-SEP-1999; 9905-0153070.  
PR 13-SEP-1999; 9905-0153758.  
PR 15-SEP-1999; 9905-0154018.  
PR 16-SEP-1999; 9905-0154039.  
PR 20-SEP-1999; 9905-0154779.  
PR 22-SEP-1999; 9905-0155139.  
PR 23-SEP-1999; 9905-0155486.  
PR 24-SEP-1999; 9905-0155659.  
PR 28-SEP-1999; 9905-0156458.  
PR 29-SEP-1999; 9905-0156596.  
PR 04-OCT-1999; 9905-0157117.  
PR 05-OCT-1999; 9905-0157753.  
PR 06-OCT-1999; 9905-0157865.  
PR 07-OCT-1999; 9905-0158029.  
PR 08-OCT-1999; 9905-0158232.  
PR 12-OCT-1999; 9905-0158369.  
PR 13-OCT-1999; 9905-0159293.  
PR 13-OCT-1999; 9905-0159294.  
PR 13-OCT-1999; 9905-0159295.  
PR 14-OCT-1999; 9905-0159329.  
PR 14-OCT-1999; 9905-0159330.  
PR 14-OCT-1999; 9905-0159331.  
PR 14-OCT-1999; 9905-0159637.  
PR 14-OCT-1999; 9905-0159638.  
PR 18-OCT-1999; 9905-0159584.  
PR 18-OCT-1999; 9905-0160741.  
PR 21-OCT-1999; 9905-0160767.  
PR 21-OCT-1999; 9905-0160770.  
PR 21-OCT-1999; 9905-0160814.  
PR 21-OCT-1999; 9905-0160815.  
PR 22-OCT-1999; 9905-0160980.  
PR 22-OCT-1999; 9905-0160981.  
PR 22-OCT-1999; 9905-0160989.  
PR 25-OCT-1999; 9905-0161404.  
PR 25-OCT-1999; 9905-0161405.  
PR 25-OCT-1999; 9905-0161359.  
PR 26-OCT-1999; 9905-0161359.  
PR 26-OCT-1999; 9905-0161360.  
PR 26-OCT-1999; 9905-0161361.  
PR 28-OCT-1999; 9905-0161920.

PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 41.7%; Score 5; DB 21; Length 142;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 5 APtQE 9  
Db 87 aptqe 91  
  
RESULT 46  
AAC08988  
ID AAC08988 standard; Protein; 148 AA.  
AC AAC08988;  
XX  
XX 17-OCT-2000 (first entry)  
DT XX  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6742.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 29-MAR-1999; 99US-0126264.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0128845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134376.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.

PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Query Match 41.7%; Score 5; DB 21; Length 142;  
Best Local Similarity 100.0%; Pred. No. 3, 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 APR0E 9  
11111  
Db 87 aptce 91  
  
RESULT 45  
AAG53843 standard; Protein; 142 AA.  
ID AAG53843  
AC AAG53843;  
XX  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 68588.  
XX  
XX  
KM Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
XX  
OS Arabidopsis thaliana.  
PN EPI033405-A2.  
XX  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX  
25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135533.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140891.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.

PR 04-MAY-1999;	9905--0132484.	PR 22-JUL-1999;	9905-0145085
PR 05-MAY-1999;	9905-0132485.	PR 22-JUL-1999;	9905-0145087
PR 06-MAY-1999;	9905-0132486.	PR 22-JUL-1999;	9905-0145089
PR 07-MAY-1999;	9905-0132487.	PR 23-JUL-1999;	9905-0145145
PR 11-MAY-1999;	9905-0132488.	PR 23-JUL-1999;	9905-0145145
PR 14-MAY-1999;	9905-0134218.	PR 26-JUL-1999;	9905-0145218
PR 14-MAY-1999;	9905-0134219.	PR 26-JUL-1999;	9905-0145276
PR 14-MAY-1999;	9905-0134221.	PR 27-JUL-1999;	9905-0145313
PR 14-MAY-1999;	9905-0134370.	PR 27-JUL-1999;	9905-0145318
PR 18-MAY-1999;	9905-0134376.	PR 27-JUL-1999;	9905-0145519
PR 19-MAY-1999;	9905-0134941.	PR 28-JUL-1999;	9905-0145521
PR 20-MAY-1999;	9905-0135121.	PR 28-JUL-1999;	9905-0145551
PR 21-MAY-1999;	9905-0135353.	PR 02-AUG-1999;	9905-0146386
PR 24-MAY-1999;	9905-0135629.	PR 02-AUG-1999;	9905-0146388
PR 25-MAY-1999;	9905-0136021.	PR 03-AUG-1999;	9905-0146389
PR 27-MAY-1999;	9905-0136392.	PR 03-AUG-1999;	9905-0147038
PR 28-MAY-1999;	9905-0136782.	PR 04-AUG-1999;	9905-0147038
PR 01-JUN-1999;	9905-0137222.	PR 05-AUG-1999;	9905-0147302
PR 03-JUN-1999;	9905-0137528.	PR 05-AUG-1999;	9905-0147191
PR 04-JUN-1999;	9905-0137502.	PR 06-AUG-1999;	9905-0147260
PR 07-JUN-1999;	9905-0137724.	PR 06-AUG-1999;	9905-0147316
PR 08-JUN-1999;	9905-0138094.	PR 06-AUG-1999;	9905-0147410
PR 10-JUN-1999;	9905-0138540.	PR 09-AUG-1999;	9905-0147493
PR 10-JUN-1999;	9905-0138847.	PR 09-AUG-1999;	9905-0147933
PR 14-JUN-1999;	9905-0139119.	PR 10-AUG-1999;	9905-0148171
PR 16-JUN-1999;	9905-0139452.	PR 11-AUG-1999;	9905-0148341
PR 16-JUN-1999;	9905-0139453.	PR 12-AUG-1999;	9905-0148341
PR 17-JUN-1999;	9905-0139492.	PR 13-AUG-1999;	9905-0148565
PR 18-JUN-1999;	9905-0139454.	PR 13-AUG-1999;	9905-0148565
PR 18-JUN-1999;	9905-0139455.	PR 16-AUG-1999;	9905-0149368
PR 18-JUN-1999;	9905-0139456.	PR 17-AUG-1999;	9905-0149175
PR 18-JUN-1999;	9905-0139457.	PR 18-AUG-1999;	9905-0149426
PR 18-JUN-1999;	9905-0139458.	PR 20-AUG-1999;	9905-0149722
PR 18-JUN-1999;	9905-0139459.	PR 20-AUG-1999;	9905-0149723
PR 18-JUN-1999;	9905-0139460.	PR 20-AUG-1999;	9905-0149929
PR 18-JUN-1999;	9905-0139461.	PR 23-AUG-1999;	9905-0149902
PR 18-JUN-1999;	9905-0139462.	PR 23-AUG-1999;	9905-0149930
PR 18-JUN-1999;	9905-0139463.	PR 25-AUG-1999;	9905-0150566
PR 18-JUN-1999;	9905-0139750.	PR 26-AUG-1999;	9905-0150884
PR 18-JUN-1999;	9905-0139763.	PR 27-AUG-1999;	9905-0151065
PR 21-JUN-1999;	9905-0139817.	PR 27-AUG-1999;	9905-0151066
PR 22-JUN-1999;	9905-0139899.	PR 27-AUG-1999;	9905-0151080
PR 23-JUN-1999;	9905-0140353.	PR 30-AUG-1999;	9905-0151303
PR 23-JUN-1999;	9905-0140354.	PR 31-AUG-1999;	9905-0151438
PR 24-JUN-1999;	9905-0140695.	PR 01-SEP-1999;	9905-0151930
PR 28-JUN-1999;	9905-0140822.	PR 07-SEP-1999;	9905-0152263
PR 29-JUN-1999;	9905-0140991.	PR 10-SEP-1999;	9905-0153070
PR 30-JUN-1999;	9905-0141287.	PR 13-SEP-1999;	9905-0153758
PR 01-JUL-1999;	9905-0141842.	PR 15-SEP-1999;	9905-0154018
PR 01-JUL-1999;	9905-0142154.	PR 16-SEP-1999;	9905-0154039
PR 02-JUL-1999;	9905-0142055.	PR 20-SEP-1999;	9905-0154779
PR 06-JUL-1999;	9905-0142390.	PR 22-SEP-1999;	9905-0155139
PR 08-JUL-1999;	9905-0142803.	PR 23-SEP-1999;	9905-0155348
PR 09-JUL-1999;	9905-0142920.	PR 24-SEP-1999;	9905-0155659
PR 12-JUL-1999;	9905-0143297.	PR 28-SEP-1999;	9905-0156458
PR 13-JUL-1999;	9905-0143562.	PR 29-SEP-1999;	9905-0156596
PR 14-JUL-1999;	9905-0143624.	PR 04-OCT-1999;	9905-0157117
PR 15-JUL-1999;	9905-0144005.	PR 05-OCT-1999;	9905-0157753
PR 16-JUL-1999;	9905-0144085.		

PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159884.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 41.7%; Score 5; DB 21; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
 |||||  
 Db 63 gevkk 67

## RESULT 43

ID AAY36760 standard; Protein; 135 AA.

XX AAY36760;

DT 07-OCT-1999 (first entry)

DE Amino acid sequence of a Chlamydia trachomatis protein.

KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;

KW paratrachoma; inclusion conjunctivitis; genital disease; peritphalitis;

KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;

XX Bartholinitis; pneumopathy; venereal lymphogranulomatosis.

OS Chlamydia trachomatis.

XX MO928475-A2.

PD 10-JUN-1999.

XX 27-NOV-1998; 98MO-IB01939.

XX 04-NOV-1998; 98US-0107077.

PR 28-NOV-1997; 97FR-0015041.

PR 17-DEC-1997; 97FR-0016034.

PA (GEST ) GENSET.  
 XX Griffais R;  
 XX WPI; 1999-371125/31.  
 XX Genome sequence of Chlamydia trachomatis  
 PS Disclosure; Page 662-663; 1755pp; English.  
 XX AAY36754-Y13949 are encoded by open reading frames (ORFs) of the genome  
 CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as  
 CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences  
 CC can also be used to control growth of the microorganism. Chlamydia  
 CC trachomatis is responsible for a large number of diseases, e.g. eye  
 CC diseases such as conventional trachoma, nonendemic trachoma,  
 CC paratrachoma, and inclusion conjunctivitis; genital diseases such as  
 CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,  
 CC peritphalitis, Bartholinitis; pneumopathy in breast feeding infants;  
 CC and venereal lymphogranulomatosis. The polypeptides of the invention  
 CC may be of use in treating these diseases.

Sequence 135 AA;

Query Match 41.7%; Score 5; DB 20; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
 |||||  
 Db 42 gevkk 46

## RESULT 44

ID AAG08989 standard; Protein; 142 AA.

XX AAG08989;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 6743.

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX Arabidopsis thaliana.

OS EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129645.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139465.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141847.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-01455913.  
PR 27-JUL-1999; 99US-01455918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 01-SEP-1999; 99US-0152363.  
PR 07-SEP-1999; 99US-0153070.  
PR 10-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 15-SEP-1999; 99US-0154038.  
PR 16-SEP-1999; 99US-0154779.  
PR 20-SEP-1999; 99US-0155139.  
PR 22-SEP-1999; 99US-0155486.  
PR 23-SEP-1999; 99US-0155659.  
PR 24-SEP-1999; 99US-0155659.



DE		Staphylococcus aureus protein of unknown function.
XX		
KW	Staphylococcus aureus protein; immune response induction; eye infection;	
KM	antibody production; T-cell immune response; gastrointestinal infection;	
KW	respiratory infection; inhibitor; bacterial infection; cardiac infection;	
KM	central nervous system; kidney infection; urinary tract infection;	
KW	antimicrobial compound identification; broad spectrum antibiotic;	
therapy.		
OS	Staphylococcus aureus.	
XX		
FH	Key Location/Qualifiers	
FT	Misc-difference 1..130	
FT	/note= "residues designated X are unspecified, and represented as xaa in the specification"	
XX		
PN	EP841394-A2.	
XX		
PD	13-MAY-1998.	
XX		
PF	24-SEP-1997; 97EP-0307485.	
XX		
FR	24-SEP-1996; 96US-0027032.	
XX		
PA	(SMIK ) SWITHKLINE BEECHAM CORP.	
PA	(SMIK ) SWITHKLINE BEECHAM PLC.	
XX		
PI	Black MT, Burnham MKR, Hodgson JE, Knowles DJC;	
PI	Lometto MA, Nicholas RO, Pratt JM, Reichard RW, Rosenberg M;	
PI	Ward JM;	
DR	WP1: 1998-252940/23.	
XX	N-PSDB; AAV53451.	
XX		
PT	New nucleic acid sequences from Staphylococcus aureus WCHU29 -	
PT	useful in vaccines and for treatment of bacterial infections of e.g.	
PT	respiratory tract and central nervous system	
XX		
PS	Claim 11: Page 312; 390pp; English.	
XX		
CC	This sequence represents a Staphylococcus aureus protein of unknown	
CC	function, and is encoded by a DNA sequence of the invention.	
CC	The DNA sequences were isolated from Staphylococcus aureus WCHU29	
CC	(NCIMB 40771). Host cells containing the DNA sequences are used to	
CC	produce polypeptides or fragments. The proteins are used in the treatment	
CC	of disease, for inducing an immune response by administering them, to	
CC	produce antibody and/or T-cell immune response. Antagonists of the	
CC	proteins are used for the inhibition of bacterial polypeptides.	
CC	Conditions which may be treated include bacterial infections, especially	
CC	respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,	
CC	urinary tract, skin, bones and joints. The proteins can also be used to	
CC	identify antimicrobial compounds which are broad spectrum antibiotics,	
CC	especially useful in the treatment of H. pylori infection.	
XX		
SO	Sequence 130 AA:	
Query Match	.41.7%; Score 5; DB 19; Length 130;	
Best Local Similarity	100.0%; Pred. No. 2.9e+02;	
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY 4 VAPTO 8		
Db 108 vaptq 112		
RESULT 41		
AAV35155		
ID AAV35155 standard; Protein; 132 AA.		
AC AAV35155;		
XX		

DT	13-SEP-1999	(first entry)
XX		
DE	Chlamydia pneumoniae transmembrane protein sequence.	
XX		
KW	Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;	
KW	sinusitis; purulent otitis media; erythema nodosum; pharyngitis;	
KW	vaccine; neutralising epitope.	
XX		
OS	Chlamydia pneumoniae.	
XX		
PN	MO9927105-A2.	
XX		
PD	03-JUN-1999.	
XX		
PF	20-NOV-1998; 98MO-IB01890.	
XX		
PR	04-NOV-1998; 98US-0107078.	
PR	21-NOV-1997; 97FR-0014673.	
XX		
PA	(GEST ) GENSET.	
XX		
PI	Griffats R;	
DR	WPI; 1999-357842/30.	
XX		
PT	Genome sequence of Chlamydia pneumoniae	
XX		
PS	Page 1021: Disclosure; 1912pp; English.	
XX		
CC	AAV34584-Y35879 represent the proteins encoded by all the open reading	
CC	frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.	
CC	C. pneumoniae causes respiratory disease such as pneumonia and	
CC	bronchitis and is thought to be a contributing factor in heart	
CC	disease, sarcoidosis, sinusitis, purulent otitis media, erythema	
CC	nodosum or pharyngitis. The polypeptides encoded by the open reading	
CC	frames of the C. pneumoniae genome (see AAV34584-Y35879) can be used in	
CC	immunogenic compositions as vaccines. Vectors containing C. pneumoniae	
CC	nucleotide sequences can also be used as immunogenic compositions,	
CC	especially where the vector directs the expression of a neutralising	
CC	epitope of C. pneumoniae.	
XX		
SO	Sequence 132 AA;	
OY	8 OEYKK 12	
DB	39 geVKK 43	
RESULT 42		
ID	AA659953	
XX	AA659953 standard; Protein; 134 AA.	
XX		
AC	AA659953;	
XX		
DT	18-OCT-2000 (first entry)	
XX		
DE	Arabidopsis thaliana protein fragment SPQ ID NO: 77605.	
XX		
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX		
OS	Arabidopsis thaliana.	
XX		
PN	EP1033405-A2.	
XX		
PD	06-SEP-2000.	
XX		

```
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148564.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150854.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151086.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 07-SEP-1999; 99US-0153070.
PR 10-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154078.
PR 15-SEP-1999; 99US-0154319.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158359.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      41.7%; Score 5; DB 21; Length 125;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 3 PVAPT 7
Db 68 pvapt 72

RESULT 40
AAW77657
ID AAW77657 standard; Protein: 130 AA.
XX AAW77657;
AC
XX
DT 30-OCT-1998 (first entry)
```



```
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154029.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 41.7%; Score 5; DB 21; Length 121;
Best Local Similarity 100.0%; Pred. No. 2,7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PVAPT 7
DB 64 pvapt 68

RESULT 38
AAB40264
ID AAB40264 standard; Protein; 125 AA.
XX
XX AAB40264;
AC
XX
XX 08-FEB-2001 (first entry)
DT
XX Human ORFX ORF28 polypeptide sequence SEQ ID NO:56.
DE
XX Human; open reading frame; ORFX; detection; cytosolic; hepatotropic;
KW vulnerability; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
```

PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149992.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156438.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159584.  
PR 18-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160771.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 21-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 41.78; Score 5; DB 21; Length 114;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYKK 12  
Db 43 qevkk 47

RESULT 37  
AAG54455  
ID AAG54455 standard; Protein: 121 AA.  
XX  
AC AAG54455;

XX 18-OCT-2000 (first entry)  
DT Zea mays protein fragment SEQ ID NO: 69435.  
XX  
DE Protein identification: signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence; corn.  
XX  
OS Zea mays subsp. mays.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132663.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 01-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.

RESULT 36  
AAG59954  
ID AAG59954 standard: Protein: 114 AA.  
XX AAG59954;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 77606.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PM EP103405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134376.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145219.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147939.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.

PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139453.	PR 12-AUG-1999;	99US-0148341.
PR 17-JUN-1999;	99US-0139452.	PR 13-AUG-1999;	99US-0148565.
PR 18-JUN-1999;	99US-0139454.	PR 14-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139455.	PR 15-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139456.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139457.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139463.	PR 25-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150084.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 21-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 22-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151067.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151080.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 28-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157753.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144335.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 21-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0145086.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160741.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 23-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 26-JUL-1999;	99US-0145224.	PR 25-OCT-1999;	99US-0161404.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161992.
PR 04-AUG-1999;	99US-0147204.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147302.	PR 29-OCT-1999;	99US-0162142.
PR 05-AUG-1999;	99US-0147192.		
PR 06-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		

Query Match

Best Local Similarity 41.7%; Score 5; DB 21; Length 114;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEKK 12	IIII
DB 43 gevkk 47	

PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161922.  
PR 28-OCT-1999; 99US-0161923.  
PR 29-OCT-1999; 99US-0162142.

Query Match 41.7%; Score 5; DB 21; Length 112;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVA 5  
| | | | |  
DB 83 stpva 87

## RESULT 34

AAV11113  
ID AAV11113 standard; Protein; 114 AA.

XX AAV11113;

DT 08-JUN-1999 (first entry)

DE H. pylori ORF hp6e10363\_30517031\_f3\_3 secreted protein.

KW Vaccine; probe; diagnostic; ORF; cell envelope protein;  
secreted protein; cytoplasmic protein; cellular protein.

OS Helicobacter pylori.

PN W09824475-A1.

PD 11-JUN-1998.

PF 05-DEC-1997; 97WO-US22104.

PR 14-JUL-1997; 97US-0891928.

PR 05-DEC-1996; 96US-0759625.

PR 25-MAR-1997; 97US-0823745.

PA (ASTR ) ASTRA AB.

PI Alm RA, Castriotta LM, Dolg PC, Kabok Z, Smith D;

DR MPI; 1998-333051/29.

DR N-PSDB; AAX30642.

PT New isolated Helicobacter pylori nucleic acids - used to develop  
products for the diagnosis, prevention and treatment of infection by  
H. pylori and other Helicobacter species

PS Claims 37, 41; Page 278-279; 339pp; English.

XX Recombinant or substantially pure preparations of H. pylori polypeptides  
are disclosed, together with the nucleic acids encoding them. In all,  
97 ORFs are shown. The proteins are variously cell envelope proteins,  
cytoplasmic proteins, secreted proteins or other cellular proteins.  
XX Vaccines containing the nucleic acids or proteins are claimed, as are  
probes containing at least 8 nucleotides from the nucleic acid  
sequences. The vaccines are useful for treating or reducing the risk of  
H. pylori infections, and the probes can be used diagnostically for  
detecting the presence of Helicobacter in a sample. The products are  
also of use in screening for compounds having the ability to interfere  
with the H. pylori life cycle or to inhibit H. pylori infection.

XX Sequence 114 AA;

Query Match 41.7%; Score 5; DB 19; Length 114;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVA 5  
| | | | |  
DB 99 stpva 103

## RESULT 35

AAG12210  
ID AAG12210 standard; Protein; 114 AA.

AC AAG12210;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 11234.

KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

OS Arabidopsis thaliana.

PN EP103405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0122548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.



PR 14-MAY-1999; 990S-0134219.  
PR 14-MAY-1999; 990S-0134221.  
PR 14-MAY-1999; 990S-0134370.  
PR 18-MAY-1999; 990S-0134768.  
PR 19-MAY-1999; 990S-0134941.  
PR 20-MAY-1999; 990S-0135124.  
PR 21-MAY-1999; 990S-0135353.  
PR 24-MAY-1999; 990S-0135629.  
PR 25-MAY-1999; 990S-0136021.  
PR 27-MAY-1999; 990S-0136392.  
PR 28-MAY-1999; 990S-0136782.  
PR 01-JUN-1999; 990S-0137222.  
PR 03-JUN-1999; 990S-0137528.  
PR 04-JUN-1999; 990S-0137502.  
PR 07-JUN-1999; 990S-0137724.  
PR 08-JUN-1999; 990S-0138094.  
PR 10-JUN-1999; 990S-0138540.  
PR 10-JUN-1999; 990S-0138847.  
PR 14-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139452.  
PR 16-JUN-1999; 990S-0139453.  
PR 17-JUN-1999; 990S-0139452.  
PR 18-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139455.  
PR 18-JUN-1999; 990S-0139456.  
PR 18-JUN-1999; 990S-0139457.  
PR 18-JUN-1999; 990S-0139458.  
PR 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 18-JUN-1999; 990S-0139753.  
PR 21-JUN-1999; 990S-0139817.  
PR 22-JUN-1999; 990S-0139899.  
PR 23-JUN-1999; 990S-0140353.  
PR 23-JUN-1999; 990S-0140354.  
PR 24-JUN-1999; 990S-0140695.  
PR 28-JUN-1999; 990S-0140823.  
PR 29-JUN-1999; 990S-0140891.  
PR 30-JUN-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141842.  
PR 01-JUL-1999; 990S-0142154.  
PR 02-JUL-1999; 990S-0142055.  
PR 06-JUL-1999; 990S-0142390.  
PR 08-JUL-1999; 990S-0142803.  
PR 09-JUL-1999; 990S-0142820.  
PR 12-JUL-1999; 990S-0142977.  
PR 13-JUL-1999; 990S-0143542.  
PR 14-JUL-1999; 990S-0143624.  
PR 15-JUL-1999; 990S-0144005.  
PR 16-JUL-1999; 990S-0144085.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 19-JUL-1999; 990S-0144335.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144632.  
PR 20-JUL-1999; 990S-0144884.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 21-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145085.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145089.  
PR 22-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 23-JUL-1999; 990S-0145224.  
  
PR 26-JUL-1999; 990S-0145276.  
PR 27-JUL-1999; 990S-0145913.  
PR 27-JUL-1999; 990S-0145918.  
PR 27-JUL-1999; 990S-0145919.  
PR 28-JUL-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.  
PR 02-AUG-1999; 990S-0146388.  
PR 02-AUG-1999; 990S-0146389.  
PR 03-AUG-1999; 990S-0147038.  
PR 04-AUG-1999; 990S-0147204.  
PR 04-AUG-1999; 990S-0147302.  
PR 05-AUG-1999; 990S-0147192.  
PR 05-AUG-1999; 990S-0147260.  
PR 06-AUG-1999; 990S-0147303.  
PR 06-AUG-1999; 990S-0147416.  
PR 09-AUG-1999; 990S-0147493.  
PR 09-AUG-1999; 990S-0147935.  
PR 10-AUG-1999; 990S-0148171.  
PR 11-AUG-1999; 990S-0148319.  
PR 12-AUG-1999; 990S-0148341.  
PR 13-AUG-1999; 990S-0148565.  
PR 13-AUG-1999; 990S-0148584.  
PR 16-AUG-1999; 990S-0149368.  
PR 17-AUG-1999; 990S-0149175.  
PR 18-AUG-1999; 990S-0149426.  
PR 20-AUG-1999; 990S-0149722.  
PR 20-AUG-1999; 990S-0149723.  
PR 20-AUG-1999; 990S-0149929.  
PR 23-AUG-1999; 990S-0149902.  
PR 23-AUG-1999; 990S-0149930.  
PR 25-AUG-1999; 990S-0150566.  
PR 26-AUG-1999; 990S-0150884.  
PR 27-AUG-1999; 990S-0151065.  
PR 27-AUG-1999; 990S-0151066.  
PR 27-AUG-1999; 990S-0151080.  
PR 30-AUG-1999; 990S-0151303.  
PR 31-AUG-1999; 990S-0151438.  
PR 01-SEP-1999; 990S-0151930.  
PR 07-SEP-1999; 990S-0152363.  
PR 10-SEP-1999; 990S-0153070.  
PR 13-SEP-1999; 990S-0153758.  
PR 13-SEP-1999; 990S-0154018.  
PR 15-SEP-1999; 990S-0154018.  
PR 16-SEP-1999; 990S-0154039.  
PR 20-SEP-1999; 990S-0154779.  
PR 22-SEP-1999; 990S-0155139.  
PR 23-SEP-1999; 990S-0155486.  
PR 24-SEP-1999; 990S-0155659.  
PR 28-SEP-1999; 990S-0156458.  
PR 29-SEP-1999; 990S-0156596.  
PR 04-OCT-1999; 990S-0157117.  
PR 05-OCT-1999; 990S-0157753.  
PR 06-OCT-1999; 990S-0157655.  
PR 07-OCT-1999; 990S-0158029.  
PR 08-OCT-1999; 990S-0158232.  
PR 12-OCT-1999; 990S-0158369.  
PR 13-OCT-1999; 990S-0159293.  
PR 13-OCT-1999; 990S-0159294.  
PR 13-OCT-1999; 990S-0159295.  
PR 14-OCT-1999; 990S-0159329.  
PR 14-OCT-1999; 990S-0159330.  
PR 14-OCT-1999; 990S-0159331.  
PR 14-OCT-1999; 990S-0159637.  
PR 14-OCT-1999; 990S-0159638.  
PR 18-OCT-1999; 990S-0159584.  
PR 21-OCT-1999; 990S-0160741.  
PR 21-OCT-1999; 990S-0160747.  
PR 21-OCT-1999; 990S-0160768.  
PR 21-OCT-1999; 990S-0160770.  
PR 21-OCT-1999; 990S-0160814.  
PR 21-OCT-1999; 990S-0160815.  
PR 22-OCT-1999; 990S-0160980.  
PR 22-OCT-1999; 990S-0160981.  
PR 22-OCT-1999; 990S-0160989.  
PR 22-OCT-1999; 990S-0160989.

```
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 41.7%; Score 5; DB 21; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 QEVKK 12
Db 39 gevkk 43

RESULT 33
AAG54393
ID AAG54393 standard; Protein; 112 AA.
AC AAG54393;
XX 18-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 69350.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 69350.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
PN EPI03405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
```

PR 14-OCT-1999; 9905-0159330.  
PR 14-OCT-1999; 9905-0159331.  
PR 14-OCT-1999; 9905-0159637.  
PR 14-OCT-1999; 9905-0159638.  
PR 18-OCT-1999; 9905-0159584.  
PR 21-OCT-1999; 9905-0160741.  
PR 21-OCT-1999; 9905-0160767.  
PR 21-OCT-1999; 9905-0160770.  
PR 21-OCT-1999; 9905-0160814.  
PR 21-OCT-1999; 9905-0160815.  
PR 22-OCT-1999; 9905-0160980.  
PR 22-OCT-1999; 9905-0160981.  
PR 22-OCT-1999; 9905-0160989.  
PR 25-OCT-1999; 9905-0161404.  
PR 25-OCT-1999; 9905-0161405.  
PR 25-OCT-1999; 9905-0161406.  
PR 26-OCT-1999; 9905-0161359.  
PR 26-OCT-1999; 9905-0161360.  
PR 26-OCT-1999; 9905-0161361.  
PR 28-OCT-1999; 9905-0161920.  
PR 28-OCT-1999; 9905-0161992.  
PR 28-OCT-1999; 9905-0161993.  
PR 29-OCT-1999; 9905-0162142.

Query Match 41.7%; Score 5; DB 21; Length 110;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 OEYK 12  
|||||  
Db 39 gevsk 43

## RESULT 32

AG59955 standard; Protein: 110 AA.

XX AAG59955:

XX 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 77607.

KM Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

XX Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 9905-0121825.

XX 09-MAR-1999; 9905-0123180.

XX 23-MAR-1999; 9905-0123588.

XX 23-MAR-1999; 9905-0123788.

XX 29-MAR-1999; 9905-0126264.

XX 01-APR-1999; 9905-0127462.

XX 06-APR-1999; 9905-0128234.

XX 08-APR-1999; 9905-0128714.

XX 16-APR-1999; 9905-0129845.

XX 19-APR-1999; 9905-0130077.

XX 21-APR-1999; 9905-0130449.

XX 23-APR-1999; 9905-0130510.

XX 23-APR-1999; 9905-0130891.

XX 28-APR-1999; 9905-0131449.

XX 30-APR-1999; 9905-0132048.

XX 30-APR-1999; 9905-0132407.

PR 04-MAY-1999; 9905-0132484.  
PR 05-MAY-1999; 9905-0132485.  
PR 06-MAY-1999; 9905-0132486.  
PR 06-MAY-1999; 9905-0132487.  
PR 07-MAY-1999; 9905-0132863.  
PR 11-MAY-1999; 9905-0134256.  
PR 14-MAY-1999; 9905-0134218.  
PR 14-MAY-1999; 9905-0134321.  
PR 14-MAY-1999; 9905-0134321.  
PR 14-MAY-1999; 9905-0134370.  
PR 18-MAY-1999; 9905-0134768.  
PR 19-MAY-1999; 9905-0134941.  
PR 20-MAY-1999; 9905-0135124.  
PR 21-MAY-1999; 9905-0135353.  
PR 24-MAY-1999; 9905-0135629.  
PR 25-MAY-1999; 9905-0136021.  
PR 27-MAY-1999; 9905-0136392.  
PR 28-MAY-1999; 9905-0136782.  
PR 01-JUN-1999; 9905-0137222.  
PR 03-JUN-1999; 9905-0137528.  
PR 04-JUN-1999; 9905-0137502.  
PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139452.  
PR 16-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139492.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.  
PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140695.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141842.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142977.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0143624.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144684.  
PR 21-JUL-1999; 9905-0144814.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.

PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142134.  
PR 02-JUL-1999; 99US-0142053.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144334.  
  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145911.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147302.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158368.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.

XX 08-FEB-2001.  
 PD  
 XX 28-JUL-2000; 2000MO-US20538.  
 PF  
 XX 30-JUL-1999; 99US-0146634.  
 PR  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PI Resnick MA, Inga A;  
 XX  
 DR WPI; 2001-123321/13.  
 XX  
 PT New isolated mutated human p53 polypeptides for inducing toxicity in a  
 PT cell, treating cancer and identifying compounds that mimic toxic or  
 PT supertransactivating mutations -  
 XX  
 PS Claim 33; Page -: 144pp; English.  
 XX  
 CC The present invention provides a number of peptides derived from the  
 CC human p53 protein and containing specific mutations. These are useful for  
 CC isolating supertransactivating and toxic p53 mutants, and for inhibiting  
 CC cell growth and treating cancer. The present sequence is a peptide  
 CC derived from the wild-type human p53 protein shown in AAB72878.  
 CC Note: The present sequence is not shown in the specification but is  
 CC derived from that given in Genbank acc. no. X02469 (see AAB72878).  
 XX  
 SO Sequence 102 AA;

Query Match 41.7%; Score 5; DB 22; Length 102;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 PVAPT 7  
 |||||  
 Db 23 pvapt 27

RESULT 30  
 AAW5784  
 ID AAW5784 standard; Protein; 107 AA.  
 AC  
 XX AAW5784;  
 XX  
 DT 18-JAN-1999 (first entry)  
 XX  
 DE Human lymphocyte surface receptor cytoplasmic domain.  
 XX  
 KW TAC1; transmembrane activator and CAML-interactor;  
 KW calcium signal-modulating cyclophilin ligand; human;  
 KW lymphocyte surface receptor; human; B-cell; B lymphocyte;  
 KW infection; cancer; rheumatoid arthritis; autoimmune disease;  
 KW glomerulonephritis; immunosuppressive; graft versus host disease;  
 KW transplant rejection; therapy; signal transduction.  
 KW  
 OS Homo sapiens.  
 XX  
 PN MO9839361-A1.  
 PD  
 XX 11-SEP-1998.  
 PF  
 XX 03-MAR-1998; 98MO-US04270.  
 PR  
 XX 03-MAR-1997; 97US-0810572.  
 PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 XX  
 PI Bram RJ, Von Bulow G;  
 XX  
 DR WPI; 1998-506346/43.  
 DR N-PSDB; AAW57329.  
 XX

PT New isolated transmembrane activator protein - used to develop  
 PT products for treating e.g. infections, cancers, autoimmune and  
 PT inflammatory conditions, transplant rejection or graft-versus-host  
 PT disease  
 PS Claim 6; Page 72; 89pp; English.  
 XX  
 CC This polypeptide comprises the C-terminal, i.e. cytoplasmic,  
 CC domain of novel human transmembrane activator and CAML-interactor  
 CC protein TAC1 (see AAW5783). TAC1 is a lymphocyte receptor protein  
 CC that is involved in the calcium activation pathway. It is normally  
 CC present in B-lymphocytes, and to a much lesser extent in immature  
 CC T-lymphocytes, and can therefore be targeted to specifically  
 CC regulate B cell responses without affecting T cell activity. The  
 CC cytoplasmic domain of TAC1 mediates signal transduction via  
 CC Ca2+-dependent and Ca2+-independent mechanisms and includes the  
 CC CAML-binding region of TAC1. Methods are claimed for identifying a  
 CC ligand for TAC1 and for identifying immunosuppressive drugs that  
 CC selectively block the action of B lymphocytes without affecting  
 CC mature T lymphocytes. TAC1 can be activated to increase immune  
 CC system activity, e.g. for treating infections or cancers. It can  
 CC be blocked to provide immunosuppression, e.g. for treating  
 CC autoimmune and inflammatory conditions such as immune complex-  
 CC induced vasculitis, glomerulonephritis, haemolytic anaemia,  
 CC myasthenia gravis, type II collagen-induced arthritis, experimental  
 CC allergic and hyperacute xenograft rejection, rheumatoid arthritis,  
 CC systemic lupus erythematosus, transplant rejection, cancer or  
 CC graft versus host disease.  
 XX  
 SO Sequence 107 AA;

Query Match 41.7%; Score 5; DB 19; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 APTQE 9  
 |||||  
 Db 53 aptqe 57

RESULT 31  
 AAG12211  
 ID AAG12211 standard; Protein; 110 AA.  
 AC  
 XX AAG12211;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 11235.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 KW  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 PD  
 XX 06-SEP-2000.  
 PF  
 XX 25-FEB-2000; 2000EP-0301439.  
 PR  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0127462.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.

PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 41.7%; Score 5; DB 21; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STRVA 5  
 DB 69 strva 73

## RESULT 28

AAV82898  
 ID AAV82898 standard; Protein; 101 AA.

XX AAV82898;

DT 24-JUL-2000 (first entry)

DE CUB domain from BMP-1 protein of mouse.

XX BPC-1; oncogene; oncogenic; cancer; prostate; bladder; antibody;  
 KW antisense; vaccine; detection; prognosis; drug screening; BMP-1;  
 mouse.  
 XX Mus musculus.  
 OS  
 XX

PN WO200009691-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 PF 10-AUG-1999; 99WO-US18250.  
 XX  
 PR 10-AUG-1998; 98US-0095982.

PA (UROG-) UROGENESYS INC.  
 PA (AFAR/) AFAR D E.  
 PA (HUBE/) HUBERT R. S.  
 PA (LEON/) LEONG K.  
 PA (LEON/) LEONG K.  
 PA (RAIT/) RAITANO A. B.  
 PA (SAFE/) SAFEFRAN D. C.  
 PA (JAKO/) JAKOBOVITS A.

PI Afar DE, Hubert RS, Leong K, Raitano AB, Safefran DC, Jakobovits A;  
 DR WPI; 2000-206006/18.

XX  
 PT New isolated BPC-1 polypeptides, useful for developing products for the  
 PT diagnosis, staging, prognosis and treatment of cancers, particularly  
 PT prostate or bladder cancer

XX  
 PS Disclosure; Figure 3; 79pp; English.

XX BPC-1 polypeptides and polynucleotides can be used for the detection  
 CC of BPC-1 polypeptides and polynucleotides in biological samples, this  
 CC is particularly useful for detecting cancers expressing BPC-1, e.g.  
 CC prostate cancer or bladder cancer. Antibodies directed against BPC-1  
 CC or antisense polynucleotides can be used for treating such cancers.  
 CC The BPC-1 polypeptides can also be used in vaccines for treating or  
 CC inhibiting the development of a cancer expressing BPC-1. The  
 CC polypeptides and polynucleotides can also be used for detection,  
 CC prognosis, drug screening and predicting susceptibility to developing  
 CC cancer. In normal human tissues BPC-1 is only expressed in certain  
 CC tissues of the brain, however, it is expressed at high levels in  
 CC prostate cancer cells and bladder cancer cells. The BPC-1 polypeptide  
 CC comprises a CUB domain which is expressed in prostate and bladder  
 CC carcinoma cells and which shows sequence similarity with CUB domains  
 CC from other known proteins. This is a CUB domain from the BMP-1  
 CC protein of a mouse. It shows 36.3% identity with the  
 CC BPC-1 CUB domain 3 in 102 residues overlap.  
 XX

SQ Sequence 101 AA;

Query Match 41.7%; Score 5; DB 21; Length 101;  
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 VAPTO 8  
 DB 23 vaptq 27

## RESULT 29

AAV72875  
 ID AAV72875 standard; Protein; 102 AA.

XX AAV72875;

DT 09-MAY-2001 (first entry)

DE Human p53 A76T/V122A mutated peptide.

XX Human; p53 mutation; mutant; muten; gene therapy; supertransactivating;  
 KW cancer.  
 XX Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200109325-A2.

XX Protein identification: signal transduction pathway; metabolic pathway;  
KW hybridisation assay: genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 99US-0139899.

PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.

KW poplar; sweetgum; teak; mahogany; bzrp; G-box binding factor;  
 KW basic helix-loop-helix zipper; LIM domain; homeobox; MADS;  
 KW homeobox zipper; LIM domain; AP2; EREBS; zinc finger domain;  
 KW type 2 Cys2His2; CCAAT box element; MYB.  
 OS Pinus radiata.  
 PN WO200053724-A2.  
 PD 14-SEP-2000.  
 XX 09-MAR-2000; 2000WO-US06112.  
 PF 11-MAR-1999; 99US-0266513.  
 PR 18-AUG-1999; 99US-0149485.  
 XX (GENE-) GENESIS RES & DEV CORP LTD.  
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.  
 XX Wood M, McGrath A, Shenk MA, Glenn M;  
 PI WPI: 2000-579369/54.  
 DR  
 XX New isolated polynucleotide encoding a plant transcription factor for  
 PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,  
 PT having modified gene expression or modified activity of a polypeptide  
 PT -  
 PS Claim 8; Page 422; 747pp; English.  
 XX The present invention relates to novel plant transcription factors from  
 CC Eucalyptus grandis or Pinus radiata. The present sequence is one such  
 CC transcription factor. The transcription factor may be used to produce a  
 CC plant having modified gene expression such as a woody plant e.g. a  
 CC eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or  
 CC to modify the activity of a polypeptide in a plant. The transcription  
 CC factors of the present invention are members from the following families  
 CC of regulatory proteins: bzrp, bzrp family of G-box binding factors, basic  
 CC helix-loop-helix zipper, homeotic/homeobox/homeobox/MADS, homeobox  
 CC zipper, LIM domain, AP2 and EREBS, zinc finger domains of type 2  
 CC Cys2His2, CCAAT box elements and MYB.  
 CC  
 XX Sequence 48 AA:  
 SQ  
 OY 1 STPVA 5  
 | | | | |  
 DB 24 stpva 28  
 Query Match 41.7%; Score 5; DB 21; Length 48;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OS Homo sapiens.  
 XX WO9924836-A1.  
 PN 20-MAY-1999.  
 XX  
 PD 04-NOV-1998; 98WO-US23435.  
 PF 17-NOV-1997; 97US-0066100.  
 PR 07-NOV-1997; 97US-0064900.  
 PR 07-NOV-1997; 97US-0064908.  
 PR 07-NOV-1997; 97US-0064911.  
 PR 07-NOV-1997; 97US-0064912.  
 PR 07-NOV-1997; 97US-0064983.  
 PR 07-NOV-1997; 97US-0064984.  
 PR 07-NOV-1997; 97US-0064985.  
 PR 07-NOV-1997; 97US-0064987.  
 PR 07-NOV-1997; 97US-0064988.  
 PR 17-NOV-1997; 97US-0066090.  
 PR 17-NOV-1997; 97US-0066094.  
 PR 17-NOV-1997; 97US-0066095.  
 PR 17-NOV-1997; 97US-0066089.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA Carter KC, Ebner R, Endress GA, Peng P, Janat F;  
 PI Kyaw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;  
 PI Ruben SM, Shi Y, Soppet DR, Wei Y;  
 XX WPI: 1999-337740/28.  
 DR N-PSDB; AAX84935.  
 XX  
 PS Claim 11; Page 350; 507pp; English.  
 XX This sequence represents a secreted human protein encoded by the gene  
 CC clone detailed in the descriptor line. The gene can be used to generate  
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc  
 CC portion (e.g. AAX84924) for increasing the stability of the fused  
 CC protein as compared to the human protein only.  
 CC The invention relates to 125 novel genes and their fragments (nucleic  
 CC acid sequences: AAX84933-X85057; amino acid sequences AAY27567-Y27933)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 125  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAX84933 for described uses).  
 CC  
 XX Sequence 87 AA:  
 SQ  
 OY 3 PVAPT 7  
 | | | | |  
 DB 60 pvapt 64  
 Query Match 41.7%; Score 5; DB 20; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



PR 12-SEP-1997; 97US-0058668.  
 PR 12-SEP-1997; 97US-0058669.  
 PR 12-SEP-1997; 97US-0058750.  
 PR 12-SEP-1997; 97US-0058971.  
 PR 12-SEP-1997; 97US-0058972.  
 PR 12-SEP-1997; 97US-0058975.  
 PR 02-OCT-1997; 97US-0060834.  
 PR 02-OCT-1997; 97US-0060841.  
 PR 02-OCT-1997; 97US-0060844.  
 PR 02-OCT-1997; 97US-0060865.  
 PR 02-OCT-1997; 97US-0061059.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Brewer LA, Ebner R, Ferrle AM, Feng P, Greene JM, Lafleur DW;  
 PI Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM, Shl Y, Young P;  
 PI Yu GJ;  
 DR WPI; 1999-080881/07.  
 DR N-PSDB; AAX04315.  
 XX  
 PT New isolated human genes and the secreted polypeptides they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders  
 PS  
 PS Claim 11; Page 266; 380pp; English.  
 XX  
 CC This sequence represents a secreted human protein encoded by the gene  
 CC clone detailed in the descriptor line. The gene can be used to generate  
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc  
 CC portion (e.g. AAX04302) for increasing the stability of the fused  
 CC protein as compared to the human protein only.  
 CC The invention relates to 86 novel genes and their fragments (nucleic  
 CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 86  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAX04311 for described uses).  
 CC  
 XX  
 SQ Sequence 27 AA;  
 QY  
 QY 3 PVAPT 7  
 QY 9 pvapt 13  
 DB  
 RESULT 24  
 ID AAY01143 standard; Protein; 45 AA.  
 XX  
 AC AAY01143;  
 XX  
 DT 18-MAY-1999 (first entry)  
 XX  
 DE Secreted protein encoded by gene 9 clone HSIDY06.  
 XX  
 KW Human; secreted protein; gene therapy; protein therapy; tissue; cancer;  
 KW tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;  
 KW developmental abnormality; foetal deficiency; Alzheimer's disease;  
 KW cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
 KW immune deficiency disease; respiratory disorder; arthritis; skeletal;  
 KW haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
 KW cardiovascular; endocrine; gastrointestinal; asthma; diagnosis.  
 XX  
 OS Homo sapiens.

XX  
 PN WO9901020-A2.  
 XX  
 PD 14-JUN-1999.  
 XX  
 PF 30-JUN-1998; 98WO-US13608.  
 XX  
 PR 12-SEP-1997; 97US-0058663.  
 PR 01-JUL-1997; 97US-0051381.  
 PR 01-JUL-1997; 97US-0051480.  
 PR 12-SEP-1997; 97US-0058598.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Carter KC, Endress GA, Feng P, Rosen CA, Ruben SM;  
 PI WPI; 1999-105683/09.  
 DR N-PSDB; AAX22119.  
 DR  
 XX  
 PT New isolated human genes and the secreted polypeptides they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, immune deficiency diseases or blood  
 PT disorders  
 PS  
 PS Claim 11; Page 146; 179pp; English.  
 XX  
 CC The invention relates to nucleic acid sequences (AAX22111 to AAX22134)  
 CC encoding human secreted proteins (AAY01135 to AAY01158). The secreted  
 CC protein gene sequences are deposited with the ATCC under deposit number  
 CC ATCC 209118. Host cells comprising recombinant vectors containing the  
 CC nucleic acid sequences are used for the recombinant production of the  
 CC secreted proteins. The polynucleotide and amino acid sequences are useful  
 CC for are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Pathological conditions can  
 CC be also diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Specific uses are described for each of the  
 CC polynucleotides, based on which tissues they are most highly expressed  
 CC in, and include developing products for the diagnosis or treatment of  
 CC cancer, tumours, developmental abnormalities and foetal deficiencies,  
 CC autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,  
 CC schizophrenia, immunological disorders, immune deficiency diseases  
 CC (AIDS), mood disorders, respiratory disorders, arthritis, asthma,  
 CC haematopoietic disorders, neural disorders, skeletal disorders,  
 CC osteoporosis, metabolic disorders, cardiovascular disorders, endocrine  
 CC disorders or gastrointestinal disorders. The polypeptides are also useful  
 CC for identifying their binding partners. The present sequence represents a  
 CC human secreted protein (see descriptor line for gene number and clone  
 CC identification).  
 CC  
 XX  
 SQ Sequence 45 AA;  
 QY  
 QY 2 TPVAP 6  
 QY 33 tpvap 37  
 DB  
 RESULT 25  
 ID AAB33087 standard; Protein; 48 AA.  
 XX  
 AC AAB33087;  
 XX  
 DT 25-JAN-2001 (first entry)  
 XX  
 DE Pinus radiata transcription factor protein sequence #214.  
 XX  
 KW Plant; transcription factor; gene expression; eucaalyptus; pine; acacia;

XX 28-JUL-2000; 2000WO-US20538.  
 PF  
 XX 30-JUL-1999; 99US-0146634.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Resnick MA, Inga A;  
 PI  
 XX MPI; 2001-123321/13.  
 DR  
 XX  
 XX  
 PT New isolated mutated human p53 polypeptides for inducing toxicity in a  
 PT cell, treating cancer and identifying compounds that mimic toxic or  
 PT supertransactivating mutations -  
 PS  
 XX Claim 6; Page -: 144pp; English.  
 CC The present invention provides a number of peptides derived from the  
 CC human p53 protein and containing specific mutations. These are useful for  
 CC isolating supertransactivating and toxic p53 mutants, and for inhibiting  
 CC cell growth and treating cancer. The present sequence is a peptide  
 CC derived from the wild-type human p53 protein shown in AAB72878.  
 CC Note: The present sequence is not shown in the specification but is  
 CC derived from that given in Genbank acc. no. X02469 (see AAB72878).  
 CC  
 SQ Sequence 11 AA:

Query Match 41.7%; Score 5; DB 22; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
 |||||  
 Db 3 pvapt 7

## RESULT 22

AAB79982  
 ID AAB79982 standard; peptide: 25 AA.

XX AAB79982;

DT 01-MAY-1996 (first entry)

XX Signal transduction regulatory tyrosine kinase SH3 binding peptide.

XX Signal transduction; src-family thymidine kinase; lipid kinase;  
 KW tyrosine kinase; adaptor; inhibition; regulation; immunodeficiency;  
 KW allergy; inflammation; cancer; autoimmune disease; Epstein-Barr virus;  
 KW immunoproliferative disease; bovine leukaemia virus; EBV; BLV;  
 KW chronic fatigue syndrome; infectious mononucleosis.

XX Homo sapiens.

OS  
 XX  
 PN WO9524915-A1.

XX 21-SEP-1995.

PD 17-MAR-1995; 95WO-US03438.

PF 17-MAR-1995; 94US-0215116.

PR 17-MAR-1994; 94US-0215116.

XX (NAJF-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.

XX MPI; 1995-336809/43.

PI Cambier JC, Clark MR, Pleiman CM;

XX Regulation of signal transduction pathways, esp. involving  
 PT src-family thymidine kinase - useful for treatment of allergic and  
 PT inflammatory responses  
 XX

PS Claim 10; Page 101; 131pp; English.

XX AAB79982-83 are signal transduction regulatory ARH1 peptides. They  
 CC bind the SH3 domain of src family tyrosine kinases. ARH1 regulatory  
 CC peptides are capable of regulating the activity of tyrosine kinases,  
 CC lipid kinases, thymidine kinases and adaptor molecules. They can be  
 CC used in a method to regulate signal transduction in such cells as  
 CC B cells, T cells, macrophages, dendritic cells and pluripotent stem  
 CC cells. The method allows regulation and treatment of allergic and  
 CC inflammatory responses, autoimmune disease, immunodeficiency and  
 CC immunoproliferative diseases, cancer and disease caused by  
 CC Epstein-Barr and bovine leukaemia viruses.  
 CC  
 SQ Sequence 25 AA:

Query Match 41.7%; Score 5; DB 16; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
 |||||  
 Db 16 tpvap 20

## RESULT 23

AAB78130  
 ID AAB78130 standard; Protein: 27 AA.

XX AAB78130;

DT 13-APR-1999 (first entry)

XX Human secreted protein encoded by gene 5 clone HRPBD36.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
 KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
 KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
 KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
 KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
 KW osteoporosis; schizophrenia; prostate; obesity; osteoclast; thymus;  
 KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

OS  
 XX  
 FH Key Location/Qualifiers

FT Misc-difference 27 /label= unknown

XX WO9856804-A1.

PD 17-DEC-1998.

PF 11-JUN-1998; 98WO-US12125.

XX 02-OCT-1997; 97US-0061060.

PR 13-JUN-1997; 97US-0049547.

PR 13-JUN-1997; 97US-0049548.

PR 13-JUN-1997; 97US-0049549.

PR 13-JUN-1997; 97US-0049550.

PR 13-JUN-1997; 97US-0049566.

PR 13-JUN-1997; 97US-0049607.

PR 13-JUN-1997; 97US-0049608.

PR 13-JUN-1997; 97US-0049609.

PR 13-JUN-1997; 97US-0049610.

PR 13-JUN-1997; 97US-0049611.

PR 13-JUN-1997; 97US-0050566.

PR 13-JUN-1997; 97US-0050901.

PR 13-JUN-1997; 97US-0052989.

PR 08-JUL-1997; 97US-0051919.

PR 18-AUG-1997; 97US-0055984.

PR 12-SEP-1997; 97US-0058665.

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVK 11  
|||||

Db 3 tgevK 7

## RESULT 19

AAR54624

ID AAR54624 standard; Protein; 11 AA.

AC AAR54624;

DT 16-JUN-1994 (first entry)

DE Listeria innocua p60 peptide epitope.

KW Antibodies; Immunoassay; conjugate.

OS Synthetic.

PN DE4318450-A.

PD 16-DEC-1993.

PF 03-JUN-1993; 93DE-4318450.

PR 11-JUN-1992; 92DE-4219111.

PR 25-NOV-1992; 92DE-4239567.

XX (MERE ) MERCK PATENT GMBH.

PI Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;

PI Hofmann G, Bubert A, Goebel W, Koehler S;

XX WPI; 1993-406956/51.

XX New primers for PCR detection of Listeria - including individual

PT species, also new peptide(s) for raising antibodies for

PT Immunochemical detection

PS Disclosure; Fig 5; 19pp; German.

CC The sequence is that of a Listeria innocua p60 peptide epitope which

CC which may be used in the prodn. of antibodies for the detection

CC of Listeria by immunoassay (partic. ELISA). It may be used as

CC part of a method that allows determination of individual Listeria

CC species.

CC

CC

SQ Sequence 11 AA;

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

SQ

KW antigen.

XX Mycobacterium tuberculosis.

OS

XX

EH Key Location/Qualifiers

FT Modified-site 1

FT /note="O-glycosylated with an alpha-mannose residue"

XX

PN W09829132-A1.

XX

PD 09-JUL-1998.

XX

PF 29-DEC-1997; 97MO-US24189.

XX

PR 31-DEC-1996; 96US-0034003.

XX

PA (UYN ) UNIV NEW YORK STATE.

XX

PI Bellisle JT, Laal S, Zolla-Pazner S;

XX

DR WPI; 1998-38787/33.

XX

PT Early detection of mycobacterial infection - by testing a biological

PT fluid sample from a subject for the presence of antibodies reactive

PT with Mycobacterium tuberculosis antigens

XX

PS Example 3; Page 71; 170pp; English.

XX

CC Sequences AAM70109, AAM70115-W70141 are N-terminal amino acid sequences

CC of glycopeptides, or non-glycosylated peptides generated by subtilisin

CC digestion of the 45 kDa MPT 32 glycoprotein. These sequences were

CC obtained by purifying the 45 kDa protein, digesting it with subtilisin,

CC and then separating the products by reversed-phase HPLC on a column.

CC This digestion yielded 26 peptides, of which five were found to be

CC glycosylated (AAM70109, AAM70115, AAM70120, AAM70126, and AAM70128). All

CC of the peptides were subjected to both Fast Atom Bombardment-mass

CC Spectrometry (FAB-MS), and automated Edman degradation to determine their

CC molecular weight and N-terminal amino acid sequence, respectively. The

CC discovery of the complete primary structure of MPT 32 glycoprotein

CC enables further work to be performed and the possibility of it being used

CC as an antigen for the early detection of mycobacterial disease.

XX

XX

SQ Sequence 11 AA;

SQ

Query Match 41.7%; Score 5; DB 19; Length 11;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
|||||

Db 1 tpevp 5

## RESULT 21

ID AAB72848 standard; Protein; 11 AA.

AC AAB72848;

DT 09-MAY-2001 (first entry)

DE Human p53 A76T mutated peptide.

KW Human; p53 mutation; mutant; mucin; gene therapy; supertansactivating;

KW cancer.

OS Homo sapiens.

OS Synthetic.

PN W0200109325-A2.

XX 08-FEB-2001.

XX

XX

CC overall electrical charge in neutral media to ensure that the  
 CC lipopeptide is hydrophilic. The peptides AAY53301-Y53549 represents  
 CC examples of peptide epitopes used to generate the lipopeptides. These are  
 CC used in therapeutic or prophylactic compositions and vaccines to induce  
 CC specific immune responses against human immunodeficiency, hepatitis B or  
 CC papilloma viruses, p53 of melanoma or the malaria parasite.

XX Sequence 9 AA;

Query Match 41.7%; Score 5; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVK 11  
 |||||  
 DB 3 tqevk 7

RESULT 17

AY40367  
 ID AAY40367 standard; Peptide; 9 AA.

AC AAY40367;

DT 19-NOV-1999 (first entry)

DE Amino acid sequence of a HIV-1 epitope.

KW Cytotoxic T cell; T lymphocyte; CD8+ epitope; T helper cell;

KW CD4+ epitope; B epitope; lipopeptide; interferon gamma; adjuvant;

KW vaccine; tumor; infection; immune response; cytokine profile;

KW acquired immune deficiency syndrome; papilloma; cancer; hepatitis;

XX autimmune disease.

XX Human immunodeficiency virus type 1.

XX FR2774687-A1.

XX 13-AUG-1999.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

XX 06-FEB-1998; 98FR-0001439.

CC (particularly of kidney, cutaneous T cells or ovary, chronic  
 CC myelogenous leukemia or mesothelioma), allergy; and autoimmune  
 CC diseases.

XX Sequence 9 AA;

Query Match 41.7%; Score 5; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TOEVK 11  
 |||||  
 DB 3 tqevk 7

RESULT 18

AY26859  
 ID AAY26859 standard; peptide; 9 AA.

AC AAY26859;

DT 14-SEP-1999 (first entry)

DE HIV-derived lipopeptide epitope #129 for mixed micelles.

KW Micelle; microaggregate; induction; immune response; lipopeptide; CTL;

KW cytotoxic T-lymphocyte; epitope; lipid; helper T-lymphocyte; HTL; HBV;

KW tetanus; toxin; vaccine; HIV; hepatitis B virus; papilloma virus; p53;

KW melanoma; Plasmodium falciparum; malaria.

XX Synthetic.

XX Human immunodeficiency virus type 1.

XX FR2771640-A1.

XX 04-JUN-1999.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

XX 03-DEC-1997; 97FR-0015246.

Query Match 41.7%; Score 5; DB 20; Length 9;

Best Local Similarity 100.0%; Pred. No. 3.4e+05;

Disclosure: Page 36; 60pp; French.

CC The invention relates to the generation of mixed micelles or  
 CC microaggregates for inducing an immune response comprising: (a) a first  
 CC lipopeptide comprising at least one CTL (cytotoxic T-lymphocyte) epitope  
 CC and at least one lipid unit; and (b) a second lipopeptide comprising at  
 CC least one HTL (helper T-lymphocyte) epitope and at least one lipid unit  
 CC different from that of the first lipopeptide. This peptide represents  
 CC an example of a lipopeptide epitope used in the invention and is derived  
 CC from a human immunodeficiency virus type 1 (HIV-1) protein. The  
 CC immunogenic lipopeptide micelles are used in vaccines, especially  
 CC against HIV, hepatitis B virus (HBV), papilloma viruses, p53, melanoma  
 CC or Plasmodium falciparum malaria.

XX Sequence 9 AA;

PT RNA  
XX Disclosure: Figure 17, 258pp: English.  
XX  
CC This invention describes a novel method for the diagnosis of a disease  
CC caused by, or associated with, an RNA molecule that has a frameshift  
CC mutation. The method is used to diagnose age-related diseases, especially  
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's  
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,  
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II  
CC and many others listed) or susceptibility to these disorders. The method  
CC allows a definitive diagnosis of Alzheimer's disease in living patients,  
CC at an early stage. It is based on the observation that disease may be  
CC caused by mutations in RNA rather than DNA. The invention describes the  
CC use of neuronal system RNA molecules, specifically proteins including  
CC beta-amyloid precursor protein (beta-APP), the microtubule associated  
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule  
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,  
CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic  
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma  
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group  
CC protein C (HMGP-C) and neuroendocrine specific protein A.  
XX  
SQ Sequence 7 AA:

Query Match 41.7%; Score 5; DB 19; Length 7;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PVAPT 7  
| | | | |  
Db 2 pvapt 6

RESULT 15  
AAW70120  
ID AAW70120 standard; peptide: 9 AA.  
XX  
AC AAW70120:  
XX  
DT 07-DEC-1998 (first entry)  
XX  
XX Peptide produced by subtilisin digestion of 45 kDa MPT 32 glycoprotein.  
DE  
XX Glycopeptide: subtilisin: reversed-phase HPLC: Edman degradation: FAB-MS.  
KM Fast Atom Bombardment-Mass Spectrometry: mycobacteria; prokaryote;  
KW antigen.  
XX  
XX Mycobacterium tuberculosis.  
OS  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1 /note="O-glycosylated with an alpha-mannose residue"  
XX  
XX W09829132-A1.  
XX  
XX 09-JUL-1998.  
XX  
XX 29-DEC-1997; 97MO-US24189.  
XX  
XX 31-DEC-1996; 96US-0034003.  
XX  
XX (UYNV ) UNIV NEW YORK STATE.  
XX  
XX Belisle JT, Laal S, Zolla-Pazner S;  
PI  
XX  
XX WPI: 1998-387787/33.  
XX  
XX Early detection of mycobacterial infection - by testing a biological  
PT fluid sample from a subject for the presence of antibodies reactive  
XX with Mycobacterium tuberculosis antigens  
XX

PS Example 3; Page 71; 170pp: English.  
XX  
XX Sequences AAW70109, AAW70115-W70141 are N-terminal amino acid sequences  
CC of glycopeptides, or non-glycosylated peptides generated by subtilisin  
CC digestion of the 45 kDa MPT 32 glycoprotein. These sequences were  
CC obtained by purifying the 45 kDa protein, digesting it with subtilisin,  
CC and then separating the products by reversed-phase HPLC on a column.  
CC This digestion yielded 26 peptides, of which five were found to be  
CC glycosylated (AAW70109, AAW70115, AAW70120, AAW70126, and AAW70128). All  
CC of the peptides were subjected to both Fast Atom Bombardment-Mass  
CC Spectrometry (FAB-MS), and automated Edman degradation to determine their  
CC molecular weight and N-terminal amino acid sequence, respectively. The  
CC discovery of the complete primary structure of MPT 32 glycoprotein  
CC enables further work to be performed and the possibility of it being used  
CC as an antigen for the early detection of mycobacterial disease.  
XX  
SQ Sequence 9 AA:

Query Match 41.7%; Score 5; DB 19; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TPVAP 6  
| | | | |  
Db 1 tpvap 5

RESULT 16  
AAV53518  
ID AAV53518 standard; Protein: 9 AA.  
XX  
XX AAV53518:  
XX  
XX 18-JAN-2000 (first entry)  
XX  
XX HIV-1 P24 protein (aa 308-316) binds HLA-Cw.  
DE  
XX  
XX Lipopeptide: epitope: cytotoxic T lymphocyte: CTL; lipid: spacer: p53;  
KM electrical charge; hydrophilicity; vaccine; immune response; HIV; HBV;  
KW human immunodeficiency virus; hepatitis B virus; papilloma virus;  
KW melanoma; malaria; parasite.  
XX  
XX Synthetic.  
OS  
XX Human immunodeficiency virus type 1.  
XX  
XX FR2776926-A1.  
XX  
XX 08-OCT-1999.  
XX  
XX 07-APR-1998; 98FR-0004323.  
XX  
XX 07-APR-1998; 98FR-0004323.  
XX  
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.  
XX  
XX (CNRS ) CNRS CENT NAT RECH SCI.  
XX  
XX (INSP ) INST PASTEUR LILLE.  
XX  
XX Le Gal FA, Guillet JG, Gabery SH, Gras MH, Melnyk O, Tartar A;  
PI  
XX  
XX WPI: 1999-583113/50.  
XX  
XX New lipopeptide containing lipid regions and two epitopes, all  
PT separated by peptide spacers that impart hydrophilicity, useful in  
XX vaccines  
XX  
XX Disclosure: Page 23; 35pp: French.  
XX  
XX The invention relates to the generation of a lipopeptide comprising at  
CC least one auxiliary T epitope, at least one cytotoxic T lymphocyte (CTL)  
CC epitope and at least one lipid residue with (1) the epitopes and lipid  
CC portion and (11) the epitopes, being separated independently by peptide  
CC spacers. These spacers comprise sequences of amino acids which carry an

```

XX      N-PSDB; AAX13994.
XX      PT      New isolated Helicobacter polynucleotides - used to develop products
XX      PT      for the diagnosis, prevention and treatment of Helicobacter
XX      PT      infections and gastrointestinal diseases
XX      PS      Claim 8; Page 301-306; 2054pp; English.
XX      CC      This sequence represents a Helicobacter pylori GPR0 protein of the
XX      CC      invention. The polypeptides can be used for preventing or treating
XX      CC      Helicobacter infections, and gastroduodenal diseases associated with
XX      CC      these infections, including acute, chronic, and atrophic gastritis, and
XX      CC      peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
XX      CC      used for the production of antibodies. The products can also be used for
XX      CC      detection and diagnosis.
XX      SQ      Sequence      1230 AA;
XX      Query Match          50.0%; Score 6; DB 19; Length 1230;
XX      Best Local Similarity 100.0%; Pred. No. 2,2e+02;
XX      Matches      6; Conservative      0; Mismatches      0; Indels      0; Gaps      0
XX      QY      7 TOEWK 12
XX      I      I I I I I
XX      Db      173 tqevkk 178
XX      RESULT 14
XX      AAY21365
XX      AAY21365 standard; Protein; 7 AA.
XX      ID

```

XX	22-JUL-1999	(first entry)
DT		
XX		
DE	Human HUPF-I mutant protein fragment 17.	
XX		
KW	Human; beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;	
KW	frameshift mutation; age-related disease; neurodegenerative disorder;	
KW	Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;	
KW	Huntington's disease; multiple sclerosis; alcoholic liver disease;	
KW	diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;	
KW	ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;	
KW	neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;	
KW	glial fibrillary acidic protein; GRAP; p53; semaphorin III; HUPF-1;	
KW	bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMG-C; NSP-A;	
KW	high mobility group protein-C; neuroendocrine specific protein A.	
XX		
OS	Synthetic.	
OS	Homo sapiens.	
XX		
PN	WO9845322-A2.	
XX		
PD	15-OCT-1998.	
XX		
FE	02-APR-1998; 98WO-IB00705.	
XX		
PR	10-APR-1997; 97US-0043163.	
XX		
PA	(UYUT-) RIJKSUNIV UTRECHT.	
PA	(ROYA-) ROYAL NETHERLANDS ACADEMIES OF ARTS & SCI.	
PA	(UYR-) UNIV ROTTERDAM ERASMUS.	
XX		
PI	Burbach JPH, Grosveld FG, Van Leeuwen FW;	
XX		
XX	WPI; 1998-609901/51.	
DR	N-PSDB; AAX75768.	
XX		
PT	Diagnosing disease by detecting frameshift mutations in RNA or	
PT	corresponding protein mutations - used to diagnose cancer and	
PT	neurological diseases, particularly Alzheimer's disease, and also	
PT	for treatment and prevention with specific ribozymes or wild-type	



```

Db      1 pcvwkk 7
      |||||
RESULT  8
AAK54615
ID      AAK54615 standard; Protein; 6 AA.
XX
AC      AAK54615;
XX
DT      16-JUN-1994 (first entry)
XX
DE      Listeria p60 peptide epitope.
XX
KW      Listeria monocytogenes; antibodies; immunoassay; conjugate.
XX
OS      Synthetic.
XX
PN      DEA318450-A.
XX
PD      16-DEC-1993.
XX
PE      03-JUN-1993; 93DE-4318450.
XX
PR      11-JUN-1992; 92DE-4219111.
PR      25-NOV-1992; 92DE-4239567.
XX
PA      (MERCK ) MERCK PATENT GMBH.
XX
PI      Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;
PI      Hofmann G, Buberl A, Goebel W, Koehler S;
XX
DR      WPI; 1993-406956/51.
XX
PT      New primers for PCR detection of Listeria - including individual
PT      species, also new peptide(s) for raising antibodies for
PT      immunochemical detection
XX
PS      Disclosure; Page 3; 19pp; German.
XX
CC      The sequence is that of a Listeria p60 peptide epitope which
CC      which may be used in the prodn. of antibodies for the detection
CC      of Listeria by immunoassay (partic. ELISA). It may be used as
CC      part of a method that allows determination of individual Listeria
CC      species, esp. L. monocytogenes.
XX
SQ      Sequence 6 AA;

Query Match      50.0%; Score 6; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 3,4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3 PVAPRQ 8
DB      1 pvpapcq 6
      |||||
RESULT  9
AAK45159
ID      AAK45159 standard; Protein; 20 AA.
XX
AC      AAK45159;
XX
DT      16-JUN-1994 (first entry)
XX
DE      Listeria p60 peptide epitope.
XX
KW      Listeria monocytogenes; antibodies; immunoassay; conjugate.
XX
OS      Synthetic.
XX
FH      Key      Location/Qualifiers

```

```

FT      Region      1..7
FT      /note= "opt. 0-7 amino acid residues"
FT      Region      14..20
FT      /note= "opt. 0-7 amino acid residues"
XX
PN      DEA318450-A.
XX
PD      16-DEC-1993.
XX
PE      03-JUN-1993; 93DE-4318450.
XX
PR      11-JUN-1992; 92DE-4219111.
PR      25-NOV-1992; 92DE-4239567.
XX
PA      (MERCK ) MERCK PATENT GMBH.
XX
PI      Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;
PI      Hofmann G, Buberl A, Goebel W, Koehler S;
XX
DR      WPI; 1993-406956/51.
XX
PT      New primers for PCR detection of Listeria - including individual
PT      species, also new peptide(s) for raising antibodies for
PT      immunochemical detection
XX
PS      Claim 3; Page 4; 19pp; German.
XX
CC      The sequence is that of a Listeria p60 peptide epitope which
CC      which may be used in the prodn. of antibodies for the detection
CC      of Listeria by immunoassay (partic. ELISA). It may be used as
CC      part of a method that allows determination of individual Listeria
CC      species, esp. L. monocytogenes.
XX
SQ      Sequence 20 AA;

Query Match      50.0%; Score 6; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3 PVAPRQ 8
DB      8 pvpapcq 13
      |||||
RESULT 10
AAK37718
ID      AAK37718 standard; Protein; 157 AA.
XX
AC      AAK37718;
XX
DT      07-OCT-1999 (first entry)
XX
DE      Amino acid sequence of a Chlamydia trachomatis protein.
XX
KW      Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW      paratrachoma; inclusion conjunctivitis; genital disease; perithelpatitis;
KW      nongonococcal urethritis; epididymitis; salpingitis;
KW      Bartholinitis; pneumonia; venereal lymphogranulomatosis.
XX
OS      Chlamydia trachomatis.
XX
PN      WO9928475-A2.
XX
PD      10-JUN-1999.
XX
PE      27-NOV-1998; 98WO-IB01939.
XX
PR      04-NOV-1998; 98US-0107077.
PR      28-NOV-1997; 97ER-0015041.
PR      17-DEC-1997; 97ER-0016034.
XX
PA      (GEST ) GENSET.

```



PF 03-JUN-1993: 93DE-4318450.  
 XX 11-JUN-1992: 92DE-4219111.  
 PR 25-NOV-1992: 92DE-4239567.  
 XX (MERE ) MERCK PATENT GMBH.  
 PI Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;  
 PI Hofmann G, Bubert A, Goebel W, Koehler S;  
 XX WPI; 1993-406956/51.  
 XX New primers for PCR detection of *Listeria* - including individual  
 PT species, also new peptide(s) for raising antibodies for  
 PT immunochemical detection  
 XX  
 PS Disclosure; Fig 2; 19pp; German.  
 XX  
 CC The sequence is that of a *Listeria* p60 peptide epitope which  
 CC which may be used in the prodn. of antibodies for the detection  
 CC of *Listeria* by immunoassay (partic. ELISA). It may be used as  
 CC part of a method that allows determination of individual *Listeria*  
 CC species, esp. *L. monocytogenes*.  
 CC  
 SQ Sequence 10 AA;  
 XX  
 OY 3 PVAPTOEVKK 12  
 Db 1 pvaptoevkk 10  
 XX  
 RESULT 6  
 AAR45168  
 ID AAR45168 standard; Protein; 9 AA.  
 AC AAR45168;  
 DT 16-JUN-1994 (first entry)  
 XX  
 DE *Listeria* p60 peptide epitope.  
 XX  
 KM *Listeria* monocytogenes; antibodies; immunoassay; conjugate.  
 XX  
 OS Synthetic.  
 XX  
 PN DE4318450-A.  
 XX  
 PD 16-DEC-1993.  
 XX  
 PF 03-JUN-1993: 93DE-4318450.  
 XX  
 PR 11-JUN-1992: 92DE-4219111.  
 PR 25-NOV-1992: 92DE-4239567.  
 XX  
 PA (MERE ) MERCK PATENT GMBH.  
 XX  
 PI Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;  
 PI Hofmann G, Bubert A, Goebel W, Koehler S;  
 XX WPI; 1993-406956/51.  
 DR  
 XX  
 XX New primers for PCR detection of *Listeria* - including individual  
 PT species, also new peptide(s) for raising antibodies for  
 PT immunochemical detection  
 XX  
 PS Disclosure; Fig 2; 19pp; German.  
 XX  
 CC The sequence is that of a *Listeria* p60 peptide epitope which

CC which may be used in the prodn. of antibodies for the detection  
 CC of *Listeria* by immunoassay (partic. ELISA). It may be used as  
 CC part of a method that allows determination of individual *Listeria*  
 CC species, esp. *L. monocytogenes*.  
 CC  
 SQ Sequence 9 AA;  
 XX  
 OY 1 STPVAPRQ 8  
 Db 2 stpvaprq 9  
 XX  
 RESULT 7  
 AAR73903  
 ID AAR73903 standard; peptide; 7 AA.  
 AC AAR73903;  
 DT 05-DEC-1995 (first entry)  
 XX  
 DE *Listeria* monocytogenes protein p60 precursor peptide 151-157.  
 XX  
 KM *Listeria* monocytogenes; protein p60 precursor; vaccine;  
 KM meningitis related homologous antigenic sequence; MRHAS; RV-1;  
 KM immunoassay; diagnosis; treatment; prophylactic; bacterial;  
 KM viral; peptide 151-157.  
 XX  
 OS *Listeria* monocytogenes.  
 XX  
 PN WO9509232-A.  
 XX  
 PD 06-APR-1995.  
 XX  
 PF 28-SEP-1994: 94MO-CN00516.  
 XX  
 PR 28-SEP-1993: 93US-0127499.  
 XX  
 PA (SHAR/) SHARMA L R.  
 PA (VALS/) VAN ALSTYNE D.  
 XX  
 PI Sharma LR, Van Alstyne D;  
 XX  
 DR WPI; 1995-147431/19.  
 XX  
 XX New peptide(s) and corresp. antibodies for the treatment of  
 PT meningitis - the peptide(s) corresp. to homologous antigenic  
 PT sites on bacterial and viral agents and on chemokine(s), used for  
 PT detecting and preventing meningitis  
 XX  
 PS Claim 35; Page 76; 98pp; English.  
 XX  
 CC AAR73913 is the *Listeria* monocytogenes protein p60 precursor. It  
 CC contains the meningitis related antigenic sequences (MRHAS) claimed  
 CC in AAR73891-R73894 and AAR73903-R73906, which are recognised by a  
 CC monoclonal antibody from the hybridoma Rubella virus (RV)-1. The  
 CC claimed MRHAS peptides may be used in immunoassays to diagnose the  
 CC presence of bacterial and/or viral meningitis agents in a sample,  
 CC or in prophylactic and therapeutic meningitis treatments. The  
 CC peptides may also be used as vaccines against meningitis.  
 CC  
 SQ Sequence 7 AA;  
 XX  
 OY 6 PTOEVKK 12  
 XX  
 Query Match 58.3%; Score 7; DB 16; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX Sequence 21 AA;  
SQ

Query Match 100.0%; Score 12; DB 16; Length 21;  
Best Local Similarity 100.0%; Pred. No. 4e-06;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVAPTOEVKK 12  
| | | | | | | | | | | | | | | | | | | | | |  
DB 3 stpvaptgevk 14

RESULT 3  
AAR45178  
ID AAR45178 standard; Protein: 478 AA.

XX AAR45178;

AC 16-JUN-1994 (first entry)

DT Listeria p60 protein.

DE Immunogenic polypeptide; antibodies; immunoassay; conjugate.

KW Listeria monocytogenes.

OS DE4318450-A.

PN 16-DEC-1993.

PD 03-JUN-1993; 93DE-4318450.

PF 11-JUN-1992; 92DE-4219111.

PR 25-NOV-1992; 92DE-4238567.

XX (MERCK) MERCK PATENT GMBH.

PI Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;

PI Hofmann G, Bubert A, Goebel W, Koehler S;

DR WPI: 1993-406956/51.

XX New primers for PCR detection of Listeria - including individual

PT species, also new peptide(s) for raising antibodies for

PT immunochemical detection

PS Disclosure: Fig 4; 19pp; German.

CC The sequence is that of the Listeria p60 protein. Antibodies

CC generated against the protein can be used in the detection

CC of Listeria by immunoassay (partic. ELISA). The detection method

CC allows determination of individual Listeria species, esp.

CC L. monocytogenes.

SQ Sequence 478 AA;

Query Match 100.0%; Score 12; DB 14; Length 478;

Best Local Similarity 100.0%; Pred. No. 7.1e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVAPTOEVKK 12

DB 146 stpvaptgevk 157

RESULT 4

AAR73913

ID AAR73913 standard; Protein: 484 AA.

XX AAR73913;

DT 05-DEC-1995 (first entry)

DE Listeria monocytogenes protein p60 precursor.

XX Listeria monocytogenes; protein p60 precursor; vaccine;

KW meningitis related homologous antigenic sequence; MRHAS; RV-1;

KW immunoassay; diagnosis; treatment; prophylactic; bacterial;

XX viral.

OS Listeria monocytogenes.

XX WO9509232-A.

XX 06-APR-1995.

XX 28-SEP-1994; 94WO-CA00516.

XX 28-SEP-1993; 93US-0127499.

XX (SHAR/) SHARMA L R.

XX (VALS/) VAN ALSTYNE D.

XX Sharma LR, Van Alstyne D;

PI WPI: 1995-147431/19.

DR New peptide(s) and corresp. antibodies for the treatment of

XX meningitis - the peptide(s) corresp. to homologous antigenic

PT sites on bacterial and viral agents and on chemokine(s), used for

PT detecting and preventing meningitis

XX Claim 47; Fig 7/10; 98pp; English.

XX AAR73913 is the Listeria monocytogenes protein p60 precursor. It

CC contains the meningitis related antigenic sequences (MRHAS) claimed

CC in AAR73891-R73894 and AAR73903-R73906, which are recognised by a

CC monoclonal antibody from the hybridoma Rubella virus (RV)-1. The

CC claimed MRHAS peptides may be used in immunoassays to diagnose the

CC presence of bacterial and/or viral meningitis agents in a sample,

CC or in prophylactic and therapeutic meningitis treatments. The

CC peptides may also be used as vaccines against meningitis.

CC NB: Identified by matching corresponding MRHAS peptides.

XX Sequence 484 AA;

OY 1 STPVAPTOEVKK 12

DB 146 stpvaptgevk 157

RESULT 5

AAR45173

ID AAR45173 standard; Protein: 10 AA.

XX AAR45173;

XX 16-JUN-1994 (first entry)

DE Listeria p60 peptide epitope.

XX Listeria monocytogenes; antibodies; immunoassay; conjugate.

XX Synthetic.

XX DE4318450-A.

XX 16-DEC-1993.

961	4	33.3	114	14	AAR53361	N-terminally trunc
962	4	33.3	114	20	AAV49136	Consensus sequence
963	4	33.3	114	20	AAV43698	Chlamydia pneumoni
964	4	33.3	114	20	AAV95454	N-terminally trunc
965	4	33.3	114	21	AAV71550	Humanised antibody
966	4	33.3	114	21	AAV03726	Human secreted pro
967	4	33.3	114	21	AAV50974	Human FVIII heavy
968	4	33.3	114	21	AAV75983	Murine skin cell p
969	4	33.3	114	22	AAV76033	Murine skin cell p
970	4	33.3	114	22	AAV55922	Skin cell protein,
971	4	33.3	114	22	AAV55972	Skin cell protein,
972	4	33.3	115	18	AAV12437	Op-1 based morphon
973	4	33.3	115	20	AAV42461	Rat granulysin prepr
974	4	33.3	115	20	AAV3287	Human phosphodiester
975	4	33.3	115	20	AAV12133	Human 5' EST seque
976	4	33.3	115	20	AAV97653	Wheat SUG1 polypep
977	4	33.3	115	20	AAV84095	Human Kabat subgro
978	4	33.3	115	21	AAV53629	Human colon cancer
979	4	33.3	115	21	AAV12169	Humanised HBV pre-
980	4	33.3	115	21	AAV12172	Humanised HBV pre-
981	4	33.3	115	21	AAV58270	Arabidopsis thaila
982	4	33.3	115	21	AAV00036	Human secreted pro
983	4	33.3	115	21	AAV00051	Human secreted pro
984	4	33.3	116	13	AAV25726	Humanised VH regio
985	4	33.3	116	14	AAV53362	N-terminally trunc
986	4	33.3	116	16	AAV6310	Human immunoglobul
987	4	33.3	116	16	AAV78728	Rmf gene product d
988	4	33.3	116	17	AAV92218	Humanised LL2 Mab
989	4	33.3	116	17	AAV92219	Humanised heavy cha
990	4	33.3	116	18	AAV27698	Variable heavy cha
991	4	33.3	116	18	AAV22418	Reshaped human AUK
992	4	33.3	116	19	AAV37812	Humanized anti-Tac
993	4	33.3	116	19	AAV58514	Protein SEQ ID NO:
994	4	33.3	116	20	AAV95453	N-terminally trunc
995	4	33.3	116	21	AAV56080	Human secreted pro
996	4	33.3	116	21	AAV18036	Eucalyptus grandis
997	4	33.3	116	21	AAV18113	Eucalyptus grandis
998	4	33.3	116	21	AAV18487	Zea mays protein f
999	4	33.3	116	21	AAV00050	Human secreted pro
1000	4	33.3	116	21	AAV50971	Human FVIII antibo

## ALIGNMENTS

RESULT 1					
AAV45172	ID	AAV45172	standard; Protein; 12 AA.		
XX	XX	XX	XX	XX	XX
AC	AAV45172;				
XX	XX	XX	XX	XX	XX
DT	16-JUN-1994	(first entry)			
XX	XX	XX	XX	XX	XX
DE	Listeria p60 peptide epitope.				
XX	XX	XX	XX	XX	XX
KW	Listeria monocytogenes; antibodies; immunoassay; conjugate.				
XX	XX	XX	XX	XX	XX
OS	Synthetic.				
XX	XX	XX	XX	XX	XX
PN	DE4318450-A.				
PD	16-DEC-1993.				
XX	XX	XX	XX	XX	XX
PF	03-JUN-1993;	93DE-4318450.			
XX	XX	XX	XX	XX	XX
PR	11-JUN-1992;	92DE-4219111.			
PR	25-NOV-1992;	92DE-4239567.			
XX	XX	XX	XX	XX	XX
PA	(MERRE ) MERCK PATENT GMBH.				
XX	XX	XX	XX	XX	XX
PI	Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;				
XX	Hofmann G, Bubert A, Goebel W, Koehler S;				
XX	XX	XX	XX	XX	XX

DR	WPI; 1993-406956/51.				
XX	XX	XX	XX	XX	XX
PT	New primers for PCR detection of Listeria - including individual species; also new peptide(s) for raising antibodies for				
PT	Immunocemical detection				
XX	XX	XX	XX	XX	XX
PS	Disclosure; Fig 2; 19pp; German.				
XX	XX	XX	XX	XX	XX
CC	The sequence is that of a Listeria p60 peptide epitope which				
CC	which may be used in the prodn. of antibodies for the detection				
CC	of Listeria by immunoassay (partic. ELISA). It may be used as				
CC	part of a method that allows determination of individual Listeria				
CC	species, esp. L. monocytogenes.				
XX	XX	XX	XX	XX	XX
SO	Sequence 12 AA:				
Query Match 100.0%; Score 12; DB 14; Length 12;					
Best Local Similarity 100.0%; Pred. No. 2.4e-06;					
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 STVPAPTOEVKK 12				
DB	1 stvpaptpgvkk 12				
RESULT 2					
AAV73891	ID	AAV73891	standard; peptide; 21 AA.		
XX	XX	XX	XX	XX	XX
AC	AAV73891;				
XX	XX	XX	XX	XX	XX
DT	05-DEC-1995	(first entry)			
XX	XX	XX	XX	XX	XX
DE	Listeria monocytogenes protein P60 precursor peptide 144-164.				
XX	XX	XX	XX	XX	XX
KW	Listeria monocytogenes; protein P60 precursor; vaccine;				
KW	meningitis related homologous antigenic sequence; MRHAS; RV-1;				
KW	immunoassay; diagnosis; treatment; prophylactic; bacterial;				
XX	XX	XX	XX	XX	XX
OS	Listeria monocytogenes.				
XX	XX	XX	XX	XX	XX
PN	W09509232-A.				
XX	XX	XX	XX	XX	XX
PD	06-APR-1995.				
XX	XX	XX	XX	XX	XX
PF	28-SEP-1994;	94WO-CA00516.			
XX	XX	XX	XX	XX	XX
PR	28-SEP-1993;	93US-0127499.			
XX	XX	XX	XX	XX	XX
PA	(SHAR/) SHARMA L.R.				
PA	(VALS/) VAN ALSTYNE D.				
XX	XX	XX	XX	XX	XX
PI	Sharma LR, Van Alstyne D;				
XX	XX	XX	XX	XX	XX
DR	WPI; 1995-147431/19.				
XX	XX	XX	XX	XX	XX
PT	New peptide(s) and corresp. antibodies for the treatment of				
PT	meningitis - the peptide(s) corresp. to homologous antigenic				
PT	sites on bacterial and viral agents and on chemokine(s), used for				
XX	XX	XX	XX	XX	XX
PS	Claim 34; Page 74; 98pp; English.				
XX	XX	XX	XX	XX	XX
CC	AAV73913 is the Listeria monocytogenes protein P60 precursor. It				
CC	contains the meningitis related antigenic sequences (MRHAS) claimed				
CC	in AAV73891-R73894 and AAV73903-R73906, which are recognised by a				
CC	monoclonal antibody from the hybridoma Rubella virus (RV)-1. The				
CC	claimed MRHAS peptides may be used in immunoassays to diagnose the				
CC	presence of bacterial and/or viral meningitis agents in a sample,				
CC	or in prophylactic and therapeutic meningitis treatments. The				
CC	peptides may also be used as vaccines against meningitis.				



669	4	33.3	73	22	AAB70025	Human CD4 D4 domai	742	4	33.3	89	21	AAG54248	Arbidopsiis thalia
670	4	33.3	73	22	AAB70026	Human CD4 D4 domai	743	4	33.3	89	21	AAG01533	Human secreted pro
671	4	33.3	73	22	AAB70027	Human CD4 D4 domai	744	4	33.3	89	21	AAG01562	Human secreted pro
672	4	33.3	73	22	AAB70028	Human CD4 D4 domai	745	4	33.3	90	21	AAG26155	Zee mays protein f
673	4	33.3	73	22	AAB70029	Human CD4 D4 domai	746	4	33.3	90	21	AAV32306	Corn acid triacylg
674	4	33.3	73	22	AAB50534	Mouse C-CAM1 cytop	747	4	33.3	91	20	AAV11416	Human 5' EST secre
675	4	33.3	74	20	AAV14475	Mouse C-CAM1 cytop	748	4	33.3	91	21	AAB42378	Human ORFX ORF2142
676	4	33.3	74	20	AAV14475	Fragment of human	749	4	33.3	91	21	AAG227140	Zee mays protein f
677	4	33.3	74	21	AAB56214	S. aureus polypept	750	4	33.3	92	20	AAV60469	Human normal blad
678	4	33.3	74	21	AAG37064	Human secreted pro	751	4	33.3	92	21	AAB40314	Human ORFX ORF78 p
679	4	33.3	74	21	AAG55678	Arbidopsiis thalia	752	4	33.3	92	21	AAG22778	Arbidopsiis thalia
680	4	33.3	74	21	AAG60223	Arbidopsiis thalia	753	4	33.3	92	21	AAG26934	Zee mays protein f
681	4	33.3	74	22	AAB50533	Human C-CAM1 cytop	754	4	33.3	92	21	AAG33234	Zee mays protein f
682	4	33.3	75	18	AAM28213	Staphylococcus aur	755	4	33.3	92	21	AAG33943	Arbidopsiis thalia
683	4	33.3	75	20	AAV30854	Human secreted pro	756	4	33.3	92	21	AAG36731	Arbidopsiis thalia
684	4	33.3	75	20	AAV35810	Chlamydia pneumoni	757	4	33.3	92	21	AAG54495	Zee mays protein f
685	4	33.3	75	20	AAV12951	Amino acid sequenc	758	4	33.3	93	19	AAV22860	SEQ ID NO. 56 from
686	4	33.3	75	21	AAG18074	Arbidopsiis thalia	759	4	33.3	93	19	AAW75090	Human secreted pro
687	4	33.3	76	20	AAV35859	Chlamydia pneumoni	760	4	33.3	93	21	AAG04405	Arbidopsiis thalia
688	4	33.3	76	20	AAV12175	Human 5' EST secre	761	4	33.3	93	21	AAG33473	Arbidopsiis thalia
689	4	33.3	76	21	AAG55739	Human secreted pro	762	4	33.3	93	21	AAG54247	Arbidopsiis thalia
690	4	33.3	76	21	AAG01671	Human secreted pro	763	4	33.3	93	21	AAG03293	Human secreted pro
691	4	33.3	76	22	AAB65860	Murine TANGO 281 c	764	4	33.3	94	15	AAR47117	Complete sequence
692	4	33.3	77	19	AAV86001	S. pneumoniae deri	765	4	33.3	94	19	AAV22861	SEQ ID NO. 57 from
693	4	33.3	77	21	AAV87219	Human secreted pro	766	4	33.3	94	19	AAW75151	Human secreted pro
694	4	33.3	78	21	AAB52108	Human secreted pro	767	4	33.3	94	19	AAW75152	Human secreted pro
695	4	33.3	78	21	AAV64691	Human 5' EST relat	768	4	33.3	94	21	AAB51698	Human secreted pro
696	4	33.3	78	22	AAB87389	Human gene 48 enco	769	4	33.3	94	21	AAG00054	Human secreted pro
697	4	33.3	78	22	AAB87412	Human gene 48 enco	770	4	33.3	94	21	AAG00065	Human secreted pro
698	4	33.3	79	19	AAV22863	SEQ ID NO. 48 enco	771	4	33.3	94	22	AAB64488	Human secreted pro
699	4	33.3	79	19	AAV22864	SEQ ID NO. 59 from	772	4	33.3	94	22	AAW20297	Human secreted pro
700	4	33.3	79	20	AAV73910	Human prostate tum	773	4	33.3	95	20	AAW74161	H. pylori cytoplas
701	4	33.3	79	21	AAB54421	Human pancreatic c	774	4	33.3	95	20	AAV60052	Human prostate tum
702	4	33.3	80	18	AAV11223	S. pneumoniae indo	775	4	33.3	95	21	AAG14413	Human endometrium
703	4	33.3	80	18	AAW27747	Staphylococcus aur	776	4	33.3	95	21	AAG57203	Arbidopsiis thalia
704	4	33.3	80	21	AAG06014	Arbidopsiis thalia	777	4	33.3	95	21	AAG59156	Arbidopsiis thalia
705	4	33.3	80	21	AAG00040	Human secreted pro	778	4	33.3	95	21	AAG00038	Human secreted pro
706	4	33.3	81	21	AAG00067	Human secreted pro	779	4	33.3	95	21	AAV91621	Human secreted pro
707	4	33.3	81	20	AAV60234	Human endometrium	780	4	33.3	96	20	AAV67917	Human secreted pro
708	4	33.3	81	20	AAV28654	Human secreted pro	781	4	33.3	96	21	AAB32768	Eucalyptus grandis
709	4	33.3	81	20	AAW89956	Antigen from Clust	782	4	33.3	96	21	AAG06013	Arbidopsiis thalia
710	4	33.3	81	21	AAV73355	Human secreted pro	783	4	33.3	96	21	AAG12504	Zee mays protein f
711	4	33.3	82	21	AAB63188	Human secreted pro	784	4	33.3	96	21	AAG15817	Arbidopsiis thalia
712	4	33.3	82	21	AAB18950	Peptide derived fr	785	4	33.3	96	21	AAG25016	Arbidopsiis thalia
713	4	33.3	82	21	AAG21847	Arbidopsiis thalia	786	4	33.3	96	21	AAG00039	Human secreted pro
714	4	33.3	82	21	AAV64976	Human 5' EST relat	787	4	33.3	96	21	AAV56644	Partial peptide fr
715	4	33.3	84	19	AAW42039	Mouse proctolista	788	4	33.3	96	21	AAV56645	Partial peptide fr
716	4	33.3	84	21	AAB03948	Human mesenchymal	789	4	33.3	96	21	AAV56647	Partial peptide fr
717	4	33.3	85	12	AAR14008	HIV-1 hxb2 gag 262	790	4	33.3	96	21	AAV56649	Partial peptide fr
718	4	33.3	85	12	AAR14458	HIV-1 hxb2 gag 262	791	4	33.3	97	14	AAV53359	Osteogenic protein
719	4	33.3	85	21	AAG02769	Human secreted pro	792	4	33.3	97	14	AAR53387	Murine osteogenic
720	4	33.3	85	21	AAV91606	Human secreted pro	793	4	33.3	97	15	AAV57475	Human HV3 VH regio
721	4	33.3	86	19	AAV38421	Proteolipid protei	794	4	33.3	97	17	AAV92078	Human HV3 VH regio
722	4	33.3	86	20	AAV73904	Human prostate tum	795	4	33.3	97	19	AAB38643	Streptococcus pneu
723	4	33.3	86	20	AAV73929	Human prostate tum	796	4	33.3	97	20	AAV72748	C. elegans CBD-6 p
724	4	33.3	86	22	AAG01952	Human secreted pro	797	4	33.3	97	20	AAW89696	Human osteogenic p
725	4	33.3	86	22	AAB35252	HIV-1 p24 protein	798	4	33.3	97	20	AAV95444	Conserved 6 cystei
726	4	33.3	87	17	AAR93003	Homologous sequenc	799	4	33.3	97	21	AAG33072	Arbidopsiis thalia
727	4	33.3	87	17	AAR92991	Homologous sequenc	800	4	33.3	97	21	AAV95681	Arbidopsiis thalia
728	4	33.3	87	21	AAG26156	Zee mays protein f	801	4	33.3	97	21	AAG00068	Human secreted pro
729	4	33.3	87	21	AAG44140	Arbidopsiis thalia	802	4	33.3	97	21	AAW90944	Comamonas testoste
730	4	33.3	87	21	AAV76169	Human secreted pro	803	4	33.3	97	21	AAV65401	Human 5' EST relat
731	4	33.3	88	13	AAR22576	Heavy chain VH16.1	804	4	33.3	97	22	AAB65852	Human TANGO 281 cy
732	4	33.3	88	21	AAG04407	Arbidopsiis thalia	805	4	33.3	98	10	AAV95681	Human osteogenic p
733	4	33.3	88	21	AAG33200	Arbidopsiis thalia	806	4	33.3	98	14	AAV33701	SP2 IgG VH segment
734	4	33.3	88	21	AAG35339	Zee mays protein f	807	4	33.3	98	14	AAR34272	Mutated human VH1
735	4	33.3	89	14	AAV34249	Arbidopsiis thalia	808	4	33.3	98	15	AAR54800	Dpi0 VH region. H
736	4	33.3	89	14	AAR34271	Dp-74 VH1 gene hea	809	4	33.3	98	16	AAV72068	Spt0 VH region. H
737	4	33.3	89	19	AAW41088	Human matrix metal	810	4	33.3	98	16	AAR72069	HV1263 VH region.
738	4	33.3	89	19	AAW41233	Alpha-v-beta-5 ant	811	4	33.3	98	19	AAV72070	OFH1.2 VH-1 H cha
739	4	33.3	89	20	AAV12949	Amino acid sequenc	812	4	33.3	98	19	AAV54008	Human Anti-CD4 ant
740	4	33.3	89	21	AAB41518	Human ORFX ORF1282	813	4	33.3	98	21	AAB40596	Human ORFX ORF460
741	4	33.3	89	21	AAG04406	Arbidopsiis thalia	814	4	33.3	98	21	AAV50949	Human anti-factor

523	4	33.3	38	15	AAR58457	TSAR binding domain	596	4	33.3	56	21	AAG01731	Human secreted pro
524	4	33.3	38	20	AAV42830	Erythropoietin receptor	597	4	33.3	56	21	AAV87131	Human TANGO 281 PS
525	4	33.3	38	20	AAV42835	Non-EPOR-binding c	598	4	33.3	56	22	AAV85854	Human TANGO 281 P
526	4	33.3	38	21	AAG55741	Arabidopsis thaliana	599	4	33.3	56	22	AAV85862	Murine TANGO 281 P
527	4	33.3	39	16	AAR66436	PCLOS 4-18 (476-49	600	4	33.3	57	21	AAV84975	Human secreted pro
528	4	33.3	39	16	AAR66410	PCLOS 4-18 (476-49	601	4	33.3	57	21	AAG54473	zeta mays protein f
529	4	33.3	39	18	AAW19025	HIV envelope glycop	602	4	33.3	58	10	AAV91360	Human secreted pro
530	4	33.3	39	20	AAV05347	HIV-1 C11VAC pep	603	4	33.3	58	10	AAG00062	Amino acids 455-51
531	4	33.3	39	20	AAV05340	HIV-1 C11VAC pep	604	4	33.3	59	16	AAV73922	Human secreted pro
532	4	33.3	39	22	AAV59899	HIV-1 C5 region of	605	4	33.3	59	16	AAV73922	B. catarrhalis CD e
533	4	33.3	40	21	AAV05912	Human endothelial	606	4	33.3	59	17	AAV04957	Partial F. anguilo
534	4	33.3	40	21	AAV05913	Bovine endothelial	607	4	33.3	59	21	AAV84976	Human secreted pro
535	4	33.3	41	19	AAV22869	SEQ ID NO. 65 from	608	4	33.3	59	21	AAV84976	Human secreted pro
536	4	33.3	41	20	AAV42648	A. thaliana D22AT3	609	4	33.3	59	21	AAV59293	Arabidopsis thaliana
537	4	33.3	41	20	AAV07893	Human secreted pro	610	4	33.3	59	21	AAV96790	Arabidopsis thaliana
538	4	33.3	42	9	AAV80686	Peptide 147 from t	611	4	33.3	59	21	AAV96790	Maize sucrose non-
539	4	33.3	42	18	AAV19055	SP410-BAL Immuno	612	4	33.3	59	22	AAV80628	Human secreted pro
540	4	33.3	42	19	AAV59179	GelA TMP-2 binding	613	4	33.3	60	18	AAV28031	Staphylococcus aur
541	4	33.3	42	20	AAV29924	Human CHD1 KRAB do	614	4	33.3	60	18	AAV15568	PRN60. Feline leu
542	4	33.3	42	20	AAV12163	Human 5' EST seque	615	4	33.3	60	19	AAV20977	Human glial fibril
543	4	33.3	43	20	AAV12848	Human 5' EST seque	616	4	33.3	60	22	AAV80561	Proline rich pepti
544	4	33.3	43	20	AAV91057	Flanking sequence	617	4	33.3	61	20	AAV27507	Human secreted pro
545	4	33.3	43	20	AAV91059	Flanking sequence	618	4	33.3	61	20	AAV27507	E. coli beta'-subu
546	4	33.3	43	20	AAW91063	Flanking sequence	619	4	33.3	61	21	AAV27507	Arabidopsis thaliana
547	4	33.3	43	20	AAW82294	Dynamlin 41.2 zein	620	4	33.3	61	21	AAV27507	Human secreted pro
548	4	33.3	43	21	AAW82294	Beta-catenin flank	621	4	33.3	62	18	AAV27824	Staphylococcus aur
549	4	33.3	43	21	AAW82294	Beta-catenin flank	622	4	33.3	62	18	AAV27824	Staphylococcus aur
550	4	33.3	43	21	AAW82294	Beta-catenin flank	623	4	33.3	62	20	AAV19648	SEQ ID NO 366 from
551	4	33.3	43	21	AAV76821	Cytochrome P450	624	4	33.3	63	16	AAV39233	Human secreted pro
552	4	33.3	44	21	AAW821848	Cytochrome P450	625	4	33.3	63	16	AAV76966	HSV-neutralising a
553	4	33.3	45	10	AAV90818	Arabidopsis thaliana	626	4	33.3	63	16	AAV76966	HSV-neutralising a
554	4	33.3	45	12	AAV87380	Non-immunogenic am	627	4	33.3	63	16	AAV76966	HSV-neutralising a
555	4	33.3	46	19	AAV79183	Human gene 39 enco	628	4	33.3	63	16	AAV76972	HSV-neutralising a
556	4	33.3	47	20	AAV36545	Fusion immunoglob	629	4	33.3	63	16	AAV76972	HSV-neutralising a
557	4	33.3	47	21	AAV83984	Fragment of human	630	4	33.3	63	16	AAV76973	HSV-neutralising a
558	4	33.3	49	19	AAW46754	Human IL-18 bindin	631	4	33.3	63	21	AAV85469	Lung cancer associ
559	4	33.3	49	20	AAV25728	Exon trap L75917.	632	4	33.3	63	21	AAV81667	Human secreted pro
560	4	33.3	49	21	AAV56202	Human secreted pro	633	4	33.3	63	21	AAV81871	zeta mays protein f
561	4	33.3	50	16	AAV71628	Human secreted pro	634	4	33.3	64	18	AAV11269	Arabidopsis thaliana
562	4	33.3	50	18	AAW29113	HIV gp120-C5. Syn	635	4	33.3	64	20	AAV18898	Type III ARP varia
563	4	33.3	51	16	AAV77798	Enhanced CCK-B/gas	636	4	33.3	64	20	AAV18898	Neisseria meningit
564	4	33.3	51	16	AAV77798	Transactivation do	637	4	33.3	64	21	AAV5463	Arabidopsis thaliana
565	4	33.3	51	18	AAW29108	Enhanced CCK-B/gas	638	4	33.3	65	19	AAV85862	S. pneumoniae derl
566	4	33.3	51	20	AAV48336	Human prostate can	639	4	33.3	65	21	AAV76042	Human secreted pro
567	4	33.3	51	20	AAV36715	Fragment of human	640	4	33.3	65	22	AAV79213	Human skin cell pr
568	4	33.3	51	21	AAV19538	Arabidopsis thaliana	641	4	33.3	65	22	AAV79213	Corynebacterium gl
569	4	33.3	51	21	AAV24626	Arabidopsis thaliana	642	4	33.3	65	22	AAV5981	Skin cell protein,
570	4	33.3	51	21	AAV34541	Arabidopsis thaliana	643	4	33.3	66	21	AAV24624	Arabidopsis thaliana
571	4	33.3	51	21	AAV34541	Arabidopsis thaliana	644	4	33.3	66	21	AAV24624	Arabidopsis thaliana
572	4	33.3	51	21	AAV34541	Arabidopsis thaliana	645	4	33.3	66	21	AAV24624	Arabidopsis thaliana
573	4	33.3	51	21	AAV34541	Arabidopsis thaliana	646	4	33.3	66	21	AAV24624	Arabidopsis thaliana
574	4	33.3	52	10	AAV91361	Hepatitis GB virus	647	4	33.3	67	19	AAV3944	Arabidopsis thaliana
575	4	33.3	52	10	AAV91361	Amino acids of the	648	4	33.3	67	19	AAV3944	Human secreted pro
576	4	33.3	52	21	AAV22970	White shrimp multi	649	4	33.3	67	20	AAV42644	Brassica napus D22
577	4	33.3	52	21	AAV66682	Membrane-bound pro	650	4	33.3	68	21	AAV53913	Human colon cancer
578	4	33.3	52	22	AAV66682	Human PRO819 (UNO4	651	4	33.3	68	21	AAV53913	Human 5' EST relat
579	4	33.3	52	22	AAV66682	Human PRO819 prote	652	4	33.3	69	18	AAV32434	Mycobacterium tube
580	4	33.3	52	22	AAV66682	Human angiotensin	653	4	33.3	69	18	AAV32434	Mycobacterium tube
581	4	33.3	53	19	AAV22898	SEQ ID NO. 94 from	654	4	33.3	69	19	AAV32434	M. tuberculosis Im
582	4	33.3	53	20	AAV30706	Amino acid sequenc	655	4	33.3	69	20	AAV39108	M. tuberculosis an
583	4	33.3	54	17	AAV05836	Mouse M-Serrate-1	656	4	33.3	69	20	AAV39108	M. tuberculosis re
584	4	33.3	54	19	AAV22867	SEQ ID NO. 63 from	657	4	33.3	70	20	AAV39108	Human 53BP2:IP-2 p
585	4	33.3	54	19	AAV22867	SEQ ID NO. 64 from	658	4	33.3	70	21	AAV39108	Human colon cancer
586	4	33.3	54	21	AAV22867	Breast and ovarian	659	4	33.3	70	21	AAV39108	Protein fragment #
587	4	33.3	54	21	AAV22867	Human secreted and	660	4	33.3	70	21	AAV39108	Arabidopsis thaliana
588	4	33.3	54	21	AAV22867	Mouse Serrate prot	661	4	33.3	71	21	AAV39108	Arabidopsis thaliana
589	4	33.3	55	8	AAV70360	Sequence of immuno	662	4	33.3	72	16	AAV71350	Consensus sequence
590	4	33.3	55	19	AAV22866	SEQ ID NO. 62 from	663	4	33.3	72	16	AAV71350	Mercury-binding re
591	4	33.3	55	19	AAV22866	SEQ ID NO. 61 from	664	4	33.3	72	16	AAV71350	Mercury-binding re
592	4	33.3	56	20	AAV26920	Relv envelope glyco	665	4	33.3	72	19	AAV22867	SEQ ID NO. 83 from
593	4	33.3	56	20	AAV26920	Clone selected aft	666	4	33.3	72	21	AAV39069	Human secreted pro
594	4	33.3	56	21	AAV54442	Human pancreatic c	667	4	33.3	72	21	AAV39069	Human secreted pro
595	4	33.3	56	21	AAV54442	Human secreted pro	668	4	33.3	73	22	AAV70018	Mammalian CD4 IL-1









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:35:42 ; Search time 56.84 Seconds  
(without alignments)  
12.799 Million cell updates/sec

Title: US-09-372-036-30

Perfect score: 12

Sequence: 1 STPVAPPTQEVKK 12

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

A.Geneseq\_0601.\*  
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	100.0	12	AA45172	Listeria p60 pepti
2	12	100.0	21	AA73891	Listeria monocytog
3	12	100.0	478	AA45178	Listeria p60 prote
4	12	100.0	484	AA73913	Listeria monocytog
5	10	83.3	10	AA45173	Listeria p60 pepti
6	8	66.7	9	AA45168	Listeria p60 pepti
7	7	58.3	7	AA73903	Listeria monocytog
8	6	50.0	6	AA54615	Listeria p60 pepti
9	6	50.0	20	AA45159	Listeria p60 pepti
10	6	50.0	157	AA73718	Amino acid sequenc
11	6	50.0	343	AA41666	Human ORFX ORF1430

12	6	50.0	750	20	AA705477	C. albicans Rb1 P
13	6	50.0	1230	19	AA988275	H. pylori GHP0 690
14	5	41.7	7	19	AA721365	Human HUPF-1 mutan
15	5	41.7	9	19	AA70120	Peptide produced b
16	5	41.7	9	20	AA73518	HIV-1 p24 protein
17	5	41.7	9	20	AA740367	Amino acid sequenc
18	5	41.7	9	20	AA726859	HIV-derived lipope
19	5	41.7	11	14	AA54624	Listeria innocua p
20	5	41.7	11	19	AA70126	Peptide produced b
21	5	41.7	11	22	AA72848	Human p53 A76T mut
22	5	41.7	25	16	AA79882	Signal transductio
23	5	41.7	27	20	AA78130	Human secreted pro
24	5	41.7	45	20	AA701143	Secreted protein e
25	5	41.7	48	21	AA73087	Pinus radiata tran
26	5	41.7	87	20	AA727569	Human secreted pro
27	5	41.7	98	21	AA54394	Arabidopsis thalia
28	5	41.7	101	21	AA782898	CUB domain from BM
29	5	41.7	102	22	AA72875	Human p53 A76T/V12
30	5	41.7	107	19	AA75784	Human lymphocyte s
31	5	41.7	110	21	AA71211	Arabidopsis thalia
32	5	41.7	110	21	AA75955	Arabidopsis thalia
33	5	41.7	112	21	AA754393	Arabidopsis thalia
34	5	41.7	114	19	AA71113	H. pylori ORF hp6e
35	5	41.7	114	21	AA71210	Arabidopsis thalia
36	5	41.7	114	21	AA75954	Arabidopsis thalia
37	5	41.7	121	21	AA754455	zeta mays protein f
38	5	41.7	125	21	AA740264	Human ORFX ORF28 p
39	5	41.7	125	21	AA754454	zeta mays protein f
40	5	41.7	130	19	AA77657	Staphylococcus aur
41	5	41.7	132	20	AA735155	Chlamydia pneumoni
42	5	41.7	134	21	AA75953	Arabidopsis thalia
43	5	41.7	135	20	AA736760	Amino acid sequenc
44	5	41.7	142	21	AA708989	Arabidopsis thalia
45	5	41.7	142	21	AA758843	Arabidopsis thalia
46	5	41.7	148	21	AA708988	Arabidopsis thalia
47	5	41.7	148	21	AA753842	Arabidopsis thalia
48	5	41.7	151	21	AA743699	Human cancer assoc
49	5	41.7	152	21	AA754453	zeta mays protein f
50	5	41.7	158	21	AA712578	zeta mays protein f
51	5	41.7	160	21	AA710791	M. avium linazaine
52	5	41.7	168	22	AA79120	Corynebacterium gl
53	5	41.7	168	22	AA79813	Corynebacterium gl
54	5	41.7	174	20	AA789904	Antigen from clust
55	5	41.7	179	21	AA755986	Arabidopsis thalia
56	5	41.7	179	21	AA759942	Arabidopsis thalia
57	5	41.7	180	21	AA755985	Arabidopsis thalia
58	5	41.7	180	21	AA759411	Arabidopsis thalia
59	5	41.7	181	21	AA725268	Arabidopsis thalia
60	5	41.7	181	21	AA708986	Arabidopsis thalia
61	5	41.7	181	21	AA710820	Arabidopsis thalia
62	5	41.7	184	22	AA780053	Corynebacterium gl
63	5	41.7	185	20	AA734781	Chlamydia pneumoni
64	5	41.7	185	21	AA754822	Arabidopsis thalia
65	5	41.7	192	16	AA769639	Hepatitis C virus
66	5	41.7	192	17	AA789510	Hepatitis C virus
67	5	41.7	193	14	AA739865	HCT27 El protein.
68	5	41.7	193	21	AA708985	Arabidopsis thalia
69	5	41.7	193	21	AA710819	Arabidopsis thalia
70	5	41.7	196	21	AA755984	Arabidopsis thalia
71	5	41.7	196	21	AA759940	Arabidopsis thalia
72	5	41.7	201	20	AA789834	Protein encoded by
73	5	41.7	202	21	AA709042	Arabidopsis thalia
74	5	41.7	202	21	AA738651	Arabidopsis thalia
75	5	41.7	205	21	AA708986	Arabidopsis thalia
76	5	41.7	205	21	AA710818	Arabidopsis thalia
77	5	41.7	214	21	AA750386	Human uncoupling p
78	5	41.7	216	22	AA75329	Human secreted pro
79	5	41.7	217	21	AA704638	Arabidopsis thalia
80	5	41.7	219	20	AA730653	Human secreted pro
81	5	41.7	221	20	AA760589	Human normal blad
82	5	41.7	226	21	AA732101	Human secreted pro
83	5	41.7	231	20	AA704036	HIV-1 (BH10 clone)
84	5	41.7	238	21	AA713699	Chlamydia sp. prot

Matches	5;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	3	QTAPK	7						
Db	217	QTAPK	221						

Search completed: August 15, 2001, 12:33:50  
Job time: 55 sec

---

NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: palin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/339,517  
FILING DATE: 14-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
FILING DATE:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 853  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-339-517-2

Query Match 45.5%; Score 5; DB 1; Length 296;  
Best local Similarity 100.0%; Pred. No. 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QTAPE 7  
DB 217 QTAPE 221

RESULT 49  
US-08-438-863-5  
Sequence 5; Application US/08438863  
Patent No. 5849585  
GENERAL INFORMATION:  
APPLICANT: Mather, Jennie P.  
APPLICANT: Ronghao Li  
APPLICANT: Chen, Jian  
TITLE OF INVENTION: ISOLATING AND CULTURING SCHWANN CELLS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: palin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,863  
FILING DATE: 10-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 946  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-438-863-5

Query Match 45.5%; Score 5; DB 2; Length 296;  
Best local Similarity 100.0%; Pred. No. 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QTAPE 7  
DB 217 QTAPE 221

RESULT 50  
US-08-438-862-5  
Sequence 5; Application US/08438862  
Patent No. 6033660  
GENERAL INFORMATION:  
APPLICANT: Mather, Jennie P.  
APPLICANT: Li, Ronghao  
APPLICANT: Chen, Jian  
TITLE OF INVENTION: ISOLATING AND CULTURING SCHWANN CELLS  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: palin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/438,862  
FILING DATE: 10-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
FILING DATE:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 946-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-438-862-5

Query Match 45.5%; Score 5; DB 3; Length 296;  
Best local Similarity 100.0%; Pred. No. 92;

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/435,436  
FILING DATE: 10-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 946-3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-435-436-5

Query Match 45.5%; Score 5; DB 1; Length 296;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QTPK 7  
Db 217 QTPK 221

RESULT 46  
US-08-428-927-2  
Sequence 2, Application US/08428927  
Patent No. 5756456  
GENERAL INFORMATION:  
APPLICANT: Ho, Wei-Hsien  
TITLE OF INVENTION: SENSOR AND MOTOR NEURON DERIVED FACTOR (SMDF)  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patln (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,927  
FILING DATE: 25-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/339517  
FILING DATE: 14-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 853D3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

US-08-428-927-2

Query Match 45.5%; Score 5; DB 1; Length 296;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QTPK 7  
Db 217 QTPK 221

RESULT 47  
US-08-428-298-2  
Sequence 2, Application US/08428298  
Patent No. 5763213  
GENERAL INFORMATION:  
APPLICANT: Ho, Wei-Hsien  
TITLE OF INVENTION: SENSOR AND MOTOR NEURON DERIVED FACTOR (SMDF)  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patln (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,298  
FILING DATE: 25-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/339517  
FILING DATE: 14-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 853D2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-428-298-2

Query Match 45.5%; Score 5; DB 1; Length 296;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QTPK 7  
Db 217 QTPK 221

RESULT 48  
US-08-339-517-2  
Sequence 2, Application US/08339517  
Patent No. 5770567  
GENERAL INFORMATION:  
APPLICANT: Ho, Wei-Hsien  
TITLE OF INVENTION: SENSOR AND MOTOR NEURON DERIVED FACTOR (SMDF)

SEQUENCE CHARACTERISTICS:  
LENGTH: 292 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US93-05000-23

Query Match 45.5%; Score 5; DB 1; Length 292;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APKAP 9  
|||||  
DB 266 APKAP 270

RESULT 43  
US-08-428-926-2  
; Sequence 2, Application US/08428926  
; Patent No. 5667780  
; GENERAL INFORMATION:  
; APPLICANT: Ho, Wei-Hsien  
; APPLICANT: Osheoff, Phyllis L.  
; TITLE OF INVENTION: SENSOR AND MOTOR NEURON DERIVED FACTOR (SMDF)  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/428,926  
FILING DATE: 25-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/339517  
FILING DATE: 14-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 853D4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-428-926-2

Query Match 45.5%; Score 5; DB 1; Length 296;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QTAPK 7  
|||||  
DB 217 QTAPK 221

RESULT 44  
US-08-435-434-5  
; Sequence 5, Application US/08435434

Patent No. 5714385  
; GENERAL INFORMATION:  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Li, Ronghao  
; APPLICANT: Chen, Jian  
; TITLE OF INVENTION: ISOLATING AND CULTURING SCHWANN CELLS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/435,434  
FILING DATE: 10-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 946-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-435-434-5

Query Match 45.5%; Score 5; DB 1; Length 296;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QTAPK 7  
|||||  
DB 217 QTAPK 221

RESULT 45  
US-08-435-436-5  
; Sequence 5, Application US/08435436  
; Patent No. 5721139  
; GENERAL INFORMATION:  
; APPLICANT: Mather, Jennie P.  
; APPLICANT: Li, Ronghao  
; APPLICANT: Chen, Jian  
; TITLE OF INVENTION: ISOLATING AND CULTURING SCHWANN CELLS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)

```

: APPLICANT: BEACH, David H.
: TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
: NUMBER OF SEQUENCES: 50
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 State Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII(text)
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/246,361A
: FILING DATE: 19-MAY-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/963,308
: FILING DATE: 16-OCT-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/888,178
: FILING DATE: 26-MAY-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/701,514
: FILING DATE: 16-MAY-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Matthew P. Vincent
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: MIT-004C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 292 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
:
: US-08-246-361A-23
:
: Query Match 45.5% Score 5: DB 2: Length 292:
: Best Local Similarity 100.0% Pred. No. 91:
: Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
:
: QY 5 APKAP 9
: |||||
: Db 266 APKAP 270
:
: RESULT 41
: US-08-463-772-23
: Sequence 23, Application US/08463772
: Patent No. 6065501
: GENERAL INFORMATION:
: APPLICANT: BEACH, David H.
: TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
: NUMBER OF SEQUENCES: 50
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: LAHIVE & COCKFIELD
: STREET: 60 State Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: ASCII(text)
:
: US-08-463-772-23
: Sequence 23, Application PC/7US9305000
: Patent No. 6065501
: GENERAL INFORMATION:
: APPLICANT: MITOTIX
: TITLE OF INVENTION: D-TYPE Cyclin and Uses Related Thereto
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Militia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: US
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/05000
: FILING DATE: 19930525
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/888,178
: FILING DATE: 26-MAY-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Granahan, Patricia
: REGISTRATION NUMBER: 32,227
: REFERENCE/DOCKET NUMBER: CSHL91-02A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-861-6240
: TELEFAX: 616-861-9540
: INFORMATION FOR SEQ ID NO: 23:
```

```

: SOFTWARE: ASCII(text)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/463,772
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/963,308
: FILING DATE: 16-OCT-1992
: APPLICATION NUMBER: US 07/888,178
: FILING DATE: 26-MAY-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/701,514
: FILING DATE: 16-MAY-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Matthew P. Vincent
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: MIT-004C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 227-5941
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 292 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
:
: US-08-463-772-23
:
: Query Match 45.5% Score 5: DB 3: Length 292:
: Best Local Similarity 100.0% Pred. No. 91:
: Matches 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
```

```

: QY 5 APKAP 9
: |||||
: Db 266 APKAP 270
:
: RESULT 42
: PCT-US93-05000-23
: Sequence 23, Application PC/7US9305000
: Patent No. 6065501
: GENERAL INFORMATION:
: APPLICANT: MITOTIX
: TITLE OF INVENTION: D-TYPE Cyclin and Uses Related Thereto
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Militia Drive
: CITY: Lexington
: STATE: Massachusetts
: COUNTRY: US
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/05000
: FILING DATE: 19930525
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/888,178
: FILING DATE: 26-MAY-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Granahan, Patricia
: REGISTRATION NUMBER: 32,227
: REFERENCE/DOCKET NUMBER: CSHL91-02A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-861-6240
: TELEFAX: 616-861-9540
: INFORMATION FOR SEQ ID NO: 23:
```

ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: CSHL91-02A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 616-861-9540  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 291 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCR-US93-05000-6

Query Match 45.5%; Score 5; DB 5; Length 291;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
|||||  
Db 265 APRAP 269

RESULT 38  
US-08-464-517-23  
Sequence 23, Application US/08464517  
Patent No. 5869640  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIT-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 292 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-464-517-23

Query Match 45.5%; Score 5; DB 2; Length 292;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
|||||  
Db 266 APRAP 270

RESULT 39  
US-08-246-361A-6  
Sequence 6, Application US/08246361A  
Patent No. 5998582  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,361A  
FILING DATE: 19-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIT-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 292 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-246-361A-6

Query Match 45.5%; Score 5; DB 2; Length 292;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
|||||  
Db 266 APRAP 270

RESULT 40  
US-08-246-361A-23  
Sequence 23, Application US/08246361A  
Patent No. 5998582  
GENERAL INFORMATION:



QY 4 TAPKA 8  
Db 92 TAPKA 96

RESULT 35  
US-08-464-517-6  
Sequence 6, Application US/08464517  
Patent No. 5869640  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIT-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEO ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 280 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-517-6

Query Match 45.5%; Score 5; DB 2; Length 280;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APKAP 9  
Db 254 APKAP 258

RESULT 36  
US-08-463-772-6  
Sequence 6, Application US/08463772  
Patent No. 6066501  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,772  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIT-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEO ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 280 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-463-772-6

Query Match 45.5%; Score 5; DB 3; Length 280;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APKAP 9  
Db 254 APKAP 258

RESULT 37  
PCT-US93-05000-6  
Sequence 6, Application PC/TUS9305000  
GENERAL INFORMATION:  
APPLICANT: MITOTIX  
TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/05000  
FILING DATE: 19930525  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/888,178  
FILING DATE: 26-MAY-1992

LENGTH: 278 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-849-420-6

Query Match 45.5%; Score 5; DB 2; Length 278;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
|||||  
DB 92 TAPKA 96

## RESULT 32

US-09-253-854-6  
Sequence 6, Application US/09253854  
Patent No. 6132717

## GENERAL INFORMATION:

APPLICANT: van de Ven, Willem Jan Marie;  
APPLICANT: van den Ouweland, Anna Maria Wilhelmina;  
APPLICANT: van Duljnhoven, Johannes Lambertus Petrus;  
APPLICANT: Robbroek, Antonius Johannes Maria; and  
APPLICANT: Konling, Plet Nico Maria  
TITLE OF INVENTION: Pharmaceutical Composition Having An  
TITLE OF INVENTION: Endoproteolytic Activity: A Process for  
TITLE OF INVENTION: Endoproteolytically Processing (Precursor)  
TITLE OF INVENTION: Proteins And For The (Micro)Biological  
TITLE OF INVENTION: Production Of Proteins  
NUMBER OF SEQUENCES: 12

## CORRESPONDENCE ADDRESSES:

ADDRESSEE: HOFFMANN & BARON, LLP  
STREET: 350 Jericho Turnpike  
CITY: Jericho  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11753

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/253,854

## CLASSIFICATION:

FILING DATE: Unassigned

## ATTORNEY/AGENT INFORMATION:

NAME: Tran, Jessica H.  
REGISTRATION NUMBER: 40,846  
REFERENCE/DOCKET NUMBER: 294-41 DIV II/CON  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
TELEX:

## INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 278 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-253-854-6

Query Match 45.5%; Score 5; DB 4; Length 278;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
|||||

DB 92 TAPKA 96

## RESULT 33

US-09-135-658-4  
Sequence 4, Application US/09135658  
Patent No. 5972683

## GENERAL INFORMATION:

APPLICANT: Tsal, Ying-Chieh  
TITLE OF INVENTION: MUTANT TYPE SUBTILISIN YAB AND ITS APPLICATION  
FILE REFERENCE: 6653-011-999  
CURRENT APPLICATION NUMBER: US/09/135,658  
CURRENT FILING DATE: 1998-08-18  
EARLIER APPLICATION NUMBER: 86112766  
EARLIER FILING DATE: 1997-09-04  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: Patentin Ver. 2.0 - beta  
SEQ ID NO 4  
LENGTH: 279  
TYPE: PRT  
ORGANISM: Bacillus subtilis  
US-09-135-658-4

Query Match 45.5%; Score 5; DB 2; Length 279;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
|||||  
DB 92 TAPKA 96

## RESULT 34

5472855-6  
Patent No. 5472855  
APPLICANT: CARTER, PAUL J.; WELLS, JAMES A.

## TITLE OF INVENTION: SUBSTRATE ASSISTED CATALYSIS

## NUMBER OF SEQUENCES: 31

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/287,964

## FILING DATE: 22-SEP-1994

## Prior Application DATA:

APPLICATION NUMBER: 90,902

FILING DATE: 12-JUL-1993

APPLICATION NUMBER: 823,039

FILING DATE: 14-JAN-1992

APPLICATION NUMBER: 35,652

FILING DATE: 06-APR-1987

APPLICATION NUMBER: 334,081

FILING DATE: 04-APR-1989

APPLICATION NUMBER: 127,134

FILING DATE: 01-DEC-1987

APPLICATION NUMBER: 846,627

FILING DATE: 01-APR-1986

APPLICATION NUMBER: 858,594

FILING DATE: 30-APR-1986

APPLICATION NUMBER: 614,612

FILING DATE: 29-MAY-1984

APPLICATION NUMBER: 614,615

FILING DATE: 29-MAY-1984

APPLICATION NUMBER: 614,617

FILING DATE: 29-MAY-1984

APPLICATION NUMBER: 614,491

FILING DATE: 29-MAY-1984

SEQ ID NO: 6;  
LENGTH: 279

5472855-6

Query Match 45.5%; Score 5; DB 6; Length 279;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
ADDRESSEE: Russell  
STREET: 200 E. Broward Boulevard  
CITY: Fort Lauderdale  
STATE: Florida  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02058  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/198,973  
FILING DATE: 18-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: C11363-16  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305/527/2498  
TELEFAX: 305/764/4396  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US95-02058-7

Query Match 45.5%; Score 5; DB 5; Length 190;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OQOTA 5  
Db 41 OQOTA 45

RESULT 30  
US-08-865-203-6  
Sequence 6, Application US/08865203  
Patent No. 5935815  
GENERAL INFORMATION:  
APPLICANT: van de Ven, Willem Jan Marie  
APPLICANT: van den Ouweland, Anna Maria Wilhelmina  
APPLICANT: van Duljnhoven, Johannes Lambertus Petrus  
APPLICANT: Robroek, Antonius Johannes Maria  
APPLICANT: Koning, Piet Nico Maria  
TITLE OF INVENTION: Pharmaceutical Composition Having An  
TITLE OF INVENTION: Endoproteolytic Activity: A Process for  
TITLE OF INVENTION: Endoproteolytically Processing (Precursor)  
TITLE OF INVENTION: Proteins And For The (Micro)Biological  
TITLE OF INVENTION: Production Of Proteins  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HOFMANN & BARON, LLP  
STREET: 350 Jericho Turnpike  
CITY: Jericho  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11753  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.24

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/865,203  
FILING DATE: 29-MAY-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Tran, Jessica H.  
REGISTRATION NUMBER: 40,846  
REFERENCE/DOCKET NUMBER: 294-41 DIV II  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 822-3550  
TELEFAX: (516) 822-3582  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 278 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-865-203-6

Query Match 45.5%; Score 5; DB 2; Length 278;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
Db 92 TAPKA 96

RESULT 31  
US-07-849-420-6  
Sequence 6, Application US/07849420  
Patent No. 5989856  
GENERAL INFORMATION:  
APPLICANT: van de Ven, Willem Jan Marie  
APPLICANT: van den Ouweland, Anna Maria Wilhelmina  
APPLICANT: van Duljnhoven, Johannes Lambertus Petrus  
APPLICANT: Robroek, Antonius Johannes Maria  
APPLICANT: Koning, Piet Nico Maria  
TITLE OF INVENTION: Pharmaceutical Composition Having An  
TITLE OF INVENTION: Endoproteolytic Activity: A Process for  
TITLE OF INVENTION: Endoproteolytically Processing (Precursor)  
TITLE OF INVENTION: Proteins And For The (Micro)Biological  
TITLE OF INVENTION: Production Of Proteins  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: COOPER & DUNHAM  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/849,420  
FILING DATE: 19920624  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Moran, Thomas F.  
REGISTRATION NUMBER: 16,579  
REFERENCE/DOCKET NUMBER: 2805/41413  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 977-9550  
TELEX: 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-487-797-7

Query Match 45.5%; Score 5; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OQOTA 5  
11111  
DB 41 OQOTA 45

## RESULT 27

US-08-701-005A-5  
Sequence 5, Application US/08701005A  
Patent No. 5877019  
GENERAL INFORMATION:  
APPLICANT: Silverman, Robert H.  
APPLICANT: Hassel, Bret A.  
APPLICANT: Zhou, Alvin  
TITLE OF INVENTION: Animal 2-5A-Dependent RNases and  
TITLE OF INVENTION: Encoding Sequence Therefor  
Patent No. 5877019  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Rothwell, F199, Ernst & Kurtz  
STREET: 555 - 13th Street, N.W., Suite 701 E  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/701,005A  
FILING DATE: 21-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/141,304  
FILING DATE: 22-OCT-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/028,086  
FILING DATE: 08-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Jondle, Robert J.  
REGISTRATION NUMBER: 33,915  
REFERENCE/DOCKET NUMBER: N1255-111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-783-6040  
TELEFAX: 202-783-6031  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: E. coli  
US-08-701-005A-5

Query Match 45.5%; Score 5; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OQOTA 5  
11111  
DB 41 OQOTA 45

## RESULT 28

US-08-479-895-5  
Sequence 5, Application US/08479895  
Patent No. 5972678  
GENERAL INFORMATION:  
APPLICANT: Silverman, Robert H.  
APPLICANT: Hassel, Bret A.  
APPLICANT: Zhou, Alvin  
TITLE OF INVENTION: Animal 2-5A-Dependent RNases and  
TITLE OF INVENTION: Encoding Sequence Therefor  
Patent No. 5972678  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
ADDRESSEE: Russell  
STREET: 200 East Broward Boulevard  
CITY: Fort Lauderdale  
STATE: Florida  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,895  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/028,086  
FILING DATE: 08-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: CL11363-13  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305/527/2498  
TELEFAX: 305/764/4996  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-479-895-5

Query Match 45.5%; Score 5; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OQOTA 5  
11111  
DB 41 OQOTA 45

## RESULT 29

PCT-US95-02058-7  
Sequence 7, Application PC/TUS9502058  
GENERAL INFORMATION:  
APPLICANT: Silverman, Robert H.  
APPLICANT: Sengupta, Dibyendu N.  
TITLE OF INVENTION: Antiviral Transgenic Plants, Vectors,  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:

Patent No. 5861300  
GENERAL INFORMATION:  
APPLICANT: Silverman, Robert H.  
APPLICANT: Sengupta, Dibyendu N.  
TITLE OF INVENTION: Antiviral Transgenic Plants, Vectors,  
TITLE OF INVENTION: Cells and Methods  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
STREET: 200 E. Broward Boulevard  
CITY: Fort Lauderdale  
STATE: Florida  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/436,771  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/198,973  
FILING DATE: 18-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: CL11363-16  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305/527/2498  
TELEFAX: 305/764/4996  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-436-771-7

Query Match 45.5%; Score 5; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOQTA 5  
|||||  
Db 41 QOQTA 45

RESULT 25  
US-08-434-998-7  
Sequence 7, Application US/08434998  
Patent No. 5866781  
GENERAL INFORMATION:  
APPLICANT: Silverman, Robert H.  
APPLICANT: Sengupta, Dibyendu N.  
TITLE OF INVENTION: Antiviral Transgenic Plants, Vectors,  
TITLE OF INVENTION: Cells and Methods  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
STREET: 200 E. Broward Boulevard  
CITY: Fort Lauderdale  
STATE: Florida  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,998  
FILING DATE:  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/198,973  
FILING DATE: 18-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: CL11363-16  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305/527/2498  
TELEFAX: 305/764/4996  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-434-998-7

Query Match 45.5%; Score 5; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOQTA 5  
|||||  
Db 41 QOQTA 45

RESULT 26  
US-08-487-797-7  
Sequence 7, Application US/08487797  
Patent No. 5866787  
GENERAL INFORMATION:  
APPLICANT: Silverman, Robert H.  
APPLICANT: Sengupta, Dibyendu N.  
TITLE OF INVENTION: Transgenic Plants Co-Expressing A  
TITLE OF INVENTION: Functional Human 2-5A System  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Holland & Knight  
STREET: One E. Broward Boulevard, #1300  
CITY: Fort Lauderdale  
STATE: Florida  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,797  
FILING DATE: 07-JUNE-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: CL11363-16(C)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305/468-7811  
TELEFAX: 305/463-2030  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid

NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454312-2074  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
STRANDEDNESS: n/a  
TOPOLOGY: linear  
MOLECULE TYPE: amino acid  
US-08-710-749-27

Query Match 45.5%; Score 5; DB 2; Length 119;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TAPKA 8  
|||||  
DB 99 TAPKA 103

RESULT 22  
US-08-905-223-464  
Sequence 464, Application US/08905223  
Patent No. 6222029  
GENERAL INFORMATION:  
APPLICANT: Edwards, Jean-Baptiste D.  
APPLICANT: Duclert, Aymeric  
APPLICANT: Lacroix, Bruno  
TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS  
NUMBER OF SEQUENCES: 503  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 501 West Broadway  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92101-3505  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Win95  
SOFTWARE: Word  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/905,223  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelson, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 464:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 135 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
TISSUE TYPE: Brain  
FEATURE:  
NAME/KEY: sig-peptide  
LOCATION: -79...-1  
IDENTIFICATION METHOD: Von Heijne matrix  
OTHER INFORMATION: score 5.  
OTHER INFORMATION: seq SFLGSAPTPIQA/LT

US-08-905-223-464

Query Match 45.5%; Score 5; DB 4; Length 135;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QTAPK 7  
|||||  
DB 8 QTAPK 12

RESULT 23  
US-08-462-481-5  
Sequence 5, Application US/08462481  
Patent No. 5840577  
GENERAL INFORMATION:  
APPLICANT: Silverman, Robert H.  
APPLICANT: Hassel, Bret A.  
APPLICANT: Zhou, Almin  
TITLE OF INVENTION: Animal 2-5A- Dependent RNases and  
TITLE OF INVENTION: Encoding Sequence Therefor  
Patent No. 5840577  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
STREET: 200 East Broward Boulevard  
CITY: Fort Lauderdale  
STATE: Florida  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,481  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/028,086  
FILING DATE: 08-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: CL1363-13  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305/527/2498  
TELEFAX: 305/764/4996  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-462-481-5

Query Match 45.5%; Score 5; DB 2; Length 190;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOOTA 5  
|||||  
DB 41 QOOTA 45

RESULT 24  
US-08-436-771-7  
Sequence 7, Application US/08436771

OY 7 KAPTE 11  
|||||  
Db 49 KAPTE 53

RESULT 19  
US-08-241-853-18  
; Sequence 18, Application US/08241853  
; Patent No. 5693488  
; GENERAL INFORMATION:  
; APPLICANT: Fang, Kathy S.  
; APPLICANT: Hanafusa, Hidesaburo  
; TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE  
; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/241,853  
; FILING DATE: 12-MAY-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-078  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 93 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; US-08-241-853-18

Query Match 45.5%; Score 5; DB 1; Length 93;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
|||||  
Db 1 APKAP 5

RESULT 20  
US-08-850-917-18  
; Sequence 18, Application US/08850917  
; Patent No. 585045  
; GENERAL INFORMATION:  
; APPLICANT: Fang, Kathy S.  
; APPLICANT: Hanafusa, Hidesaburo  
; TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE  
; TITLE OF INVENTION: AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue

CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/850,917  
; FILING DATE: 02-MAY-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/241,853  
; FILING DATE: 12-MAY-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-078  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 93 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; US-08-850-917-18

Query Match 45.5%; Score 5; DB 2; Length 93;  
Best Local Similarity 100.0%; Pred. No. 34;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
|||||  
Db 1 APKAP 5

RESULT 21  
US-08-710-749-27  
; Sequence 27, Application US/08710749  
; Patent No. 5955089  
; GENERAL INFORMATION:  
; APPLICANT: Briles, David E.  
; APPLICANT: Hollingshead, Susan  
; APPLICANT: Becker, Robert  
; TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Curtis, Morris & Safford  
; STREET: 530 Fifth Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/710,749  
; FILING DATE: 20-SEP-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/053,197A  
FILING DATE: 01-APR-1998  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MacKnight, Kamrin T.  
REGISTRATION NUMBER: 38,230  
REFERENCE/DOCKET NUMBER: UALB-03293  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 41 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-053-197A-46

Query Match 45.5%; Score 5; DB 3; Length 41;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOOTA 5  
11111  
DB 35 QOOTA 39

RESULT 17  
US-08-102-385G-15  
Sequence 15, Application US/08102385G  
Patent No. 5962294  
GENERAL INFORMATION:  
APPLICANT: Paulson, James C.  
APPLICANT: Wen, Xiaohong  
APPLICANT: Livingston, Brian Duane  
APPLICANT: Gillespie, William  
APPLICANT: Kelm, Sorge  
APPLICANT: Burlingame, Alma L.  
APPLICANT: Medzhiradzky, Katalin  
TITLE OF INVENTION: Compositions and Methods for the  
TITLE OF INVENTION: Identification and Synthesis of Stablytransferases  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Oppenheimer Wolff & Donnelly LLP  
STREET: 2029 Century Park East, 38th Floor  
CITY: Los Angeles  
STATE: CA  
COUNTRY: USA  
ZIP: 90067  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/102,385G  
FILING DATE: 04-AUG-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/925369  
FILING DATE: 04-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Oldenkamp, David J.  
REGISTRATION NUMBER: 29,421  
REFERENCE/DOCKET NUMBER: 97-062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (310)788-5000

TELEFAX: (310) 277-1297  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 55 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-102-385G-15

Query Match 45.5%; Score 5; DB 2; Length 55;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 KAPTE 11  
11111  
DB 33 KAPTE 37

RESULT 18  
US-08-102-385G-29  
Sequence 29, Application US/08102385G  
Patent No. 5962294  
GENERAL INFORMATION:  
APPLICANT: Paulson, James C.  
APPLICANT: Wen, Xiaohong  
APPLICANT: Livingston, Brian Duane  
APPLICANT: Gillespie, William  
APPLICANT: Kelm, Sorge  
APPLICANT: Burlingame, Alma L.  
APPLICANT: Medzhiradzky, Katalin  
TITLE OF INVENTION: Compositions and Methods for the  
TITLE OF INVENTION: Identification and Synthesis of Stablytransferases  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Oppenheimer Wolff & Donnelly LLP  
STREET: 2029 Century Park East, 38th Floor  
CITY: Los Angeles  
STATE: CA  
COUNTRY: USA  
ZIP: 90067  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/102,385G  
FILING DATE: 04-AUG-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/925369  
FILING DATE: 04-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Oldenkamp, David J.  
REGISTRATION NUMBER: 29,421  
REFERENCE/DOCKET NUMBER: 97-062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (310)788-5000  
TELEFAX: (310) 277-1297  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 77 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-102-385G-29

Query Match 45.5%; Score 5; DB 2; Length 77;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria innocua  
US-08-456-670B-36

Query Match 45.5%; Score 5; DB 2; Length 13;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QOTAP 6  
|||||  
DB 8 QOTAP 12

RESULT 15  
US-08-456-670B-23  
Sequence 23, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINNWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUEBEL, ANDREAS  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 25 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1-7  
OTHER INFORMATION: /note- "SOME OR ALL XAA AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 19-25  
OTHER INFORMATION: /note- "SOME OR ALL XAA AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
US-08-456-670B-23

Query Match 45.5%; Score 5; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QOTAP 6  
|||||  
DB 14 QOTAP 18

RESULT 16  
US-09-053-197A-46  
Sequence 46, Application US/09053197A  
Patent No. 6022952  
GENERAL INFORMATION:  
APPLICANT: Weiner, Joel H.  
APPLICANT: Turner, Raymond J.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN  
SECRETION  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Medien & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

```

; Patent No. 5656452
; GENERAL INFORMATION:
; APPLICANT: Rao, Anjana
; APPLICANT: Hogan, Patrick Gerald
; APPLICANT: Mc Caffrey, Patricia
; APPLICANT: Jain, Jugnu
; TITLE OF INVENTION: NF-ATP, A T LYMPHOCYTE
; TITLE OF INVENTION: DNA-BINDING PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,006C
; FILING DATE: October 29, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/017,052
; FILING DATE: February 11, 1993
; APPLICATION NUMBER: 08/006,067
; FILING DATE: January 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 890
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
; US-08-145-006C-5

```

```

Query Match          54.5%; Score 6; DB 1; Length 890;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QOOTAP 6
    111111
DB 695 QOOTAP 700

```

```

RESULT 13
; PCT-US94-00545-5
; Sequence 5, Application PC/TUS9400545
; GENERAL INFORMATION:
; APPLICANT: Rao, Anjana
; APPLICANT: Hogan, Patrick Gerald
; APPLICANT: Mc Caffrey, Patricia
; APPLICANT: Jain, Jugnu
; TITLE OF INVENTION: NF-ATP, A T LYMPHOCYTE
; TITLE OF INVENTION: DNA-BINDING PROTEIN
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.

```

```

; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00545
; FILING DATE: 18-JAN-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/145,006
; FILING DATE: October 29, 1993
; APPLICATION NUMBER: 08/017,052
; FILING DATE: February 11, 1993
; APPLICATION NUMBER: 08/006,067
; FILING DATE: January 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 890
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
; PCT-US94-00545-5

```

```

Query Match          54.5%; Score 6; DB 5; Length 890;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QOOTAP 6
    111111
DB 695 QOOTAP 700

```

```

RESULT 14
; US-08-456-670B-36
; Sequence 36, Application US/08456670B
; Patent No. 5932415
; GENERAL INFORMATION:
; APPLICANT: SCHUBERT, PETER
; APPLICANT: NEUMANN, SIEGFRIED
; APPLICANT: PAMELZIK, MARTINA
; APPLICANT: LINKWIELER, WINFRIED
; APPLICANT: BURGER, CHRISTA
; APPLICANT: HOFMANN, GOTTFRIED
; APPLICANT: BUBERT, ANDREAS
; APPLICANT: GOEBEL, WERNER
; APPLICANT: KOHLER, STERAN
; TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING
; TITLE OF INVENTION: LISTERIAS
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD., SUITE 1400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: US
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1-7  
OTHER INFORMATION: /note- "SOME OR ALL Xaa AMINO ACIDS"  
OTHER INFORMATION: MAY BE ABSENT"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 17-23  
OTHER INFORMATION: /note- "SOME OR ALL Xaa AMINO ACIDS"  
OTHER INFORMATION: MAY BE ABSENT"  
US-08-456-670B-20

Query Match 81.8%; Score 9; DB 2; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.00043;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QQTAPKAPT 10  
Db 8 QQTAPKAPT 16

RESULT 10  
US-08-127-499A-34  
Sequence 34, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
TITLE OF INVENTION: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid

STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-34

Query Match 63.6%; Score 7; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QQTAPKA 8  
Db 1 QQTAPKA 7

RESULT 11  
US-08-482-847-34  
Sequence 34, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
TITLE OF INVENTION: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-34

Query Match 63.6%; Score 7; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QQTAPKA 8  
Db 1 QQTAPKA 7

RESULT 12  
US-08-145-006C-5  
Sequence 5, Application US/08145006C

NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-26

Query Match 100.0%; Score 11; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 3.8e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 000TAPKAPTE 11  
|||||

DB 291 000TAPKAPTE 301

RESULT 8  
US-08-482-847-26  
Sequence 26, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-26

Query Match 100.0%; Score 11; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 3.8e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 000TAPKAPTE 11  
|||||

DB 291 000TAPKAPTE 301

RESULT 9  
US-08-456-670B-20  
Sequence 20, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWILLER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STERAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 421911.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302

APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 232 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-39

Query Match 100.0%; Score 11; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 2e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
Db 222 000TAPKAPTE 232

RESULT 6  
US-08-456-670B-40  
Sequence 40 Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, STEFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA

APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 478 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-40

Query Match 100.0%; Score 11; DB 2; Length 478;  
Best Local Similarity 100.0%; Pred. No. 3.8e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
Db 291 000TAPKAPTE 301

RESULT 7  
US-08-127-499A-26  
Sequence 26 Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES

APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-456-670B-42

Query Match 100.0%; Score 11; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOOTAPKAPTE 11  
Db 2 QOOTAPKAPTE 12

RESULT 3  
US-08-127-499A-33  
Sequence 33, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-33

Query Match 100.0%; Score 11; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOOTAPKAPTE 11  
Db 7 QOOTAPKAPTE 17

RESULT 4  
US-08-482-847-33  
Sequence 33, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-33

Query Match 100.0%; Score 11; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.6e-06;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOOTAPKAPTE 11  
Db 7 QOOTAPKAPTE 17

RESULT 5  
US-08-456-670B-39  
Sequence 39, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED

977 4 36.4 1457 2 US-08-087-244A-1 Sequence 1, Appl1  
978 4 36.4 1457 2 US-08-991-258A-3 Sequence 3, Appl1  
979 4 36.4 1457 2 US-08-769-399-3 Sequence 3, Appl1  
980 4 36.4 1457 3 US-08-665-259-27 Sequence 27, Appl  
981 4 36.4 1457 3 US-08-762-500-27 Sequence 27, Appl  
982 4 36.4 1457 3 US-08-991-953A-3 Sequence 3, Appl1  
983 4 36.4 1501 2 US-08-447-464-3 Sequence 3, Appl1  
984 4 36.4 1501 2 US-08-716-679-3 Sequence 3, Appl1  
985 4 36.4 1535 3 US-08-755-587-185 Sequence 185, App  
986 4 36.4 1580 2 US-08-804-227C-11 Sequence 11, Appl  
987 4 36.4 1580 2 US-08-804-198-5 Sequence 5, Appl  
988 4 36.4 1596 3 US-09-356-952-3 Sequence 3, Appl1  
989 4 36.4 1612 1 US-08-169-927-2 Sequence 2, Appl1  
990 4 36.4 1620 1 US-08-542-363-2 Sequence 2, Appl1  
991 4 36.4 1620 1 US-09-100-089-2 Sequence 1, Appl1  
992 4 36.4 1663 2 US-08-793-126-1 Sequence 1, Appl1  
993 4 36.4 1663 4 US-09-132-371-1 Sequence 1, Appl1  
994 4 36.4 1663 4 US-09-142-334-22 Sequence 22, Appl  
995 4 36.4 1664 2 US-08-642-846-2 Sequence 12, Appl  
996 4 36.4 1720 2 US-08-477-451-12 Sequence 12, Appl  
997 4 36.4 1754 1 US-07-745-206A-13 Sequence 13, Appl  
998 4 36.4 1848 4 US-08-311-363-13 Sequence 6, Appl1  
999 4 36.4 1848 4 US-08-286-791-6 Sequence 6, Appl1  
1000 4 36.4 1848 5 PCT-US95-10661A-6

## ALIGNMENTS

RESULT 1  
US-08-456-670B-29  
Sequence 29, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINXWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4

FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-29

Query Match 100.0%; Score 11; DB 2; Length 11;  
Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOOTAPKAPTE 11  
Db 1 QOOTAPKAPTE 11

RESULT 2  
US-08-456-670B-42  
Sequence 42, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINXWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4

831	4	36.4	851	3	US-08-948-547-2	Sequence 2, Appl1	904	4	36.4	921	1	US-08-188-582-2	Sequence 2, Appl1
832	4	36.4	851	4	US-09-087-465-4	Sequence 4, Appl1	905	4	36.4	921	1	US-08-396-4778-2	Sequence 2, Appl1
833	4	36.4	851	4	US-09-364-970-2	Sequence 2, Appl1	906	4	36.4	921	1	US-08-646-715-2	Sequence 2, Appl1
834	4	36.4	851	5	PCT-US95-17025-2	Sequence 2, Appl1	907	4	36.4	921	1	US-08-818-822-2	Sequence 2, Appl1
835	4	36.4	852	1	US-08-276-099A-13	Sequence 13, Appl	908	4	36.4	921	1	US-08-938-365-4	Sequence 2, Appl1
836	4	36.4	852	1	US-08-81-890-13	Sequence 13, Appl	909	4	36.4	941	2	US-08-343-760A-2	Sequence 2, Appl1
837	4	36.4	852	1	US-08-471-033-36	Sequence 36, Appl	910	4	36.4	942	1	US-08-141-324-14	Sequence 14, Appl
838	4	36.4	852	2	US-08-471-044-36	Sequence 36, Appl	911	4	36.4	942	1	US-08-541-902-14	Sequence 14, Appl
839	4	36.4	852	2	US-08-463-483A-36	Sequence 36, Appl	912	4	36.4	946	5	PCT-US95-08493-13	Sequence 13, Appl
840	4	36.4	852	2	US-08-471-046A-36	Sequence 36, Appl	913	4	36.4	968	1	US-08-426-236-2	Sequence 2, Appl1
841	4	36.4	852	2	US-08-470-566B-36	Sequence 36, Appl	914	4	36.4	979	1	US-08-308-881-6	Sequence 6, Appl1
842	4	36.4	852	2	US-08-408-519-5	Sequence 5, Appl1	915	4	36.4	979	2	US-09-058-263-6	Sequence 6, Appl1
843	4	36.4	852	2	US-08-469-334-36	Sequence 36, Appl	916	4	36.4	979	2	US-09-059-099-6	Sequence 6, Appl1
844	4	36.4	852	3	PCT-US95-03552-5	Sequence 36, Appl	917	4	36.4	979	3	US-09-058-264-6	Sequence 6, Appl1
845	4	36.4	852	3	PCT-US95-03552-5	Sequence 5, Appl1	918	4	36.4	979	5	PCT-US95-06530-6	Sequence 6, Appl1
846	4	36.4	855	2	US-08-816-693A-2	Sequence 2, Appl1	919	4	36.4	984	1	US-08-242-932-2	Sequence 2, Appl1
847	4	36.4	855	3	US-08-085-291-2	Sequence 2, Appl1	920	4	36.4	984	1	US-08-714-481-2	Sequence 2, Appl1
848	4	36.4	857	2	US-08-779-113-2	Sequence 2, Appl1	921	4	36.4	984	1	PCT-US95-06111-2	Sequence 2, Appl1
849	4	36.4	858	1	US-07-712-284-2	Sequence 2, Appl1	922	4	36.4	989	4	US-09-110-517-4	Sequence 4, Appl1
850	4	36.4	858	2	US-08-583-562B-2	Sequence 2, Appl1	923	4	36.4	992	1	US-08-127-499A-1	Sequence 1, Appl1
851	4	36.4	858	5	PCT-US92-04237-2	Sequence 2, Appl1	924	4	36.4	992	1	US-08-482-847-1	Sequence 1, Appl1
852	4	36.4	859	2	US-08-395-580-2	Sequence 2, Appl1	925	4	36.4	992	1	US-08-482-847-1	Sequence 1, Appl1
853	4	36.4	859	5	PCT-US95-02792-2	Sequence 2, Appl1	926	4	36.4	995	5	PCT-US95-04910-14	Sequence 14, Appl
854	4	36.4	860	5	PCT-US95-08493-19	Sequence 2, Appl1	927	4	36.4	1004	4	US-08-916-352-2	Sequence 2, Appl
855	4	36.4	860	5	US-08-405-392-17	Sequence 19, Appl	928	4	36.4	1018	1	US-08-408-093-6	Sequence 6, Appl1
856	4	36.4	866	3	US-08-487-691-17	Sequence 17, Appl	929	4	36.4	1018	1	US-08-408-420A-6	Sequence 6, Appl1
857	4	36.4	868	1	US-08-374-834-1	Sequence 1, Appl1	930	4	36.4	1018	1	US-08-714-901-6	Sequence 6, Appl1
858	4	36.4	868	5	US-08-644-271-1	Sequence 1, Appl1	931	4	36.4	1018	1	US-08-452-052-2	Sequence 2, Appl1
859	4	36.4	868	5	PCT-US95-08493-21	Sequence 21, Appl	932	4	36.4	1018	3	US-08-040-741-6	Sequence 6, Appl1
860	4	36.4	869	1	US-08-188-582-32	Sequence 32, Appl	933	4	36.4	1026	2	US-08-530-792D-23	Sequence 23, Appl
861	4	36.4	869	1	US-08-646-715-32	Sequence 32, Appl	934	4	36.4	1026	2	US-08-542-003-6	Sequence 6, Appl1
862	4	36.4	869	1	US-08-374-834-16	Sequence 16, Appl	935	4	36.4	1026	2	US-08-322-760A-6	Sequence 6, Appl1
863	4	36.4	869	2	US-08-644-271-29	Sequence 29, Appl	936	4	36.4	1026	2	US-08-530-792D-22	Sequence 22, Appl
864	4	36.4	872	1	US-08-766-014-2	Sequence 2, Appl1	937	4	36.4	1063	1	US-08-093-453B-3	Sequence 3, Appl1
865	4	36.4	881	4	US-08-960-780-32	Sequence 32, Appl	938	4	36.4	1063	1	US-08-127-499A-8	Sequence 8, Appl1
866	4	36.4	881	4	US-09-073-898-32	Sequence 32, Appl	939	4	36.4	1078	3	US-08-480-847-8	Sequence 8, Appl1
867	4	36.4	884	1	US-08-471-044-5	Sequence 5, Appl1	940	4	36.4	1129	3	US-08-904-871-11	Sequence 11, Appl
868	4	36.4	884	2	US-08-463-483A-5	Sequence 5, Appl1	941	4	36.4	1141	1	US-08-131-365B-4	Sequence 54, Appl
869	4	36.4	884	2	US-08-471-046A-5	Sequence 5, Appl1	942	4	36.4	1141	1	US-08-363-300-2	Sequence 2, Appl1
870	4	36.4	884	2	US-08-470-566B-5	Sequence 5, Appl1	943	4	36.4	1141	2	US-08-668-123-54	Sequence 54, Appl
871	4	36.4	884	2	US-08-469-334-5	Sequence 5, Appl1	944	4	36.4	1164	2	US-08-589-736-1	Sequence 1, Appl
872	4	36.4	884	2	US-09-300-529-5	Sequence 5, Appl1	945	4	36.4	1240	2	US-08-680-326-37	Sequence 37, Appl
873	4	36.4	885	3	US-08-310-912A-2	Sequence 2, Appl1	946	4	36.4	1248	2	US-08-348-365-17	Sequence 17, Appl
874	4	36.4	885	2	US-08-841-089-2	Sequence 2, Appl1	947	4	36.4	1248	2	US-08-465-965-17	Sequence 17, Appl
875	4	36.4	885	5	PCT-US95-04570-2	Sequence 2, Appl1	948	4	36.4	1248	3	US-08-465-966-17	Sequence 17, Appl
876	4	36.4	885	5	PCT-US95-04589-2	Sequence 2, Appl1	949	4	36.4	1257	4	US-09-220-641-3	Sequence 3, Appl1
877	4	36.4	885	5	PCT-US95-04589-2	Sequence 2, Appl1	950	4	36.4	1261	3	US-09-080-885-2	Sequence 2, Appl1
878	4	36.4	894	2	US-08-867-941-15	Sequence 15, Appl	951	4	36.4	1285	4	US-08-974-549A-600	Sequence 600, App
879	4	36.4	898	1	US-08-465-995A-2	Sequence 2, Appl1	952	4	36.4	1285	3	US-08-762-428A-6	Sequence 6, Appl1
880	4	36.4	898	1	US-08-465-995A-4	Sequence 4, Appl1	953	4	36.4	1324	2	US-08-811-897A-56	Sequence 56, Appl1
881	4	36.4	898	2	US-08-465-994C-2	Sequence 2, Appl1	954	4	36.4	1336	2	US-08-551-355-6	Sequence 6, Appl1
882	4	36.4	898	2	US-08-465-994C-4	Sequence 4, Appl1	955	4	36.4	1336	5	PCT-US93-12687-6	Sequence 6, Appl1
883	4	36.4	898	2	US-08-966-145-2	Sequence 2, Appl1	956	4	36.4	1338	1	US-08-471-033-50	Sequence 50, Appl
884	4	36.4	898	2	US-08-966-145-4	Sequence 4, Appl1	957	4	36.4	1338	2	US-08-471-044-50	Sequence 50, Appl
885	4	36.4	903	2	US-08-867-941-11	Sequence 11, Appl	958	4	36.4	1338	2	US-08-463-483A-50	Sequence 50, Appl
886	4	36.4	903	2	US-08-853-310-2	Sequence 2, Appl1	959	4	36.4	1338	2	US-08-471-046A-50	Sequence 50, Appl
887	4	36.4	905	3	US-08-405-392-2	Sequence 2, Appl1	960	4	36.4	1338	2	US-08-470-566B-50	Sequence 50, Appl
888	4	36.4	905	3	US-08-487-691-2	Sequence 2, Appl1	961	4	36.4	1338	2	US-08-469-334-50	Sequence 50, Appl
889	4	36.4	905	4	US-08-666-221B-4	Sequence 4, Appl1	962	4	36.4	1338	3	US-09-300-529-50	Sequence 50, Appl
890	4	36.4	905	4	US-08-666-221B-10	Sequence 10, Appl	963	4	36.4	1346	1	US-08-471-044-23	Sequence 23, Appl
891	4	36.4	907	3	US-08-783-774-2	Sequence 7, Appl1	964	4	36.4	1346	2	US-08-471-044-23	Sequence 23, Appl
892	4	36.4	907	3	US-08-930-996A-7	Sequence 7, Appl1	965	4	36.4	1346	2	US-08-463-483A-23	Sequence 23, Appl
893	4	36.4	907	5	PCT-US95-04611A-19	Sequence 19, Appl	966	4	36.4	1346	2	US-08-471-046A-23	Sequence 23, Appl
894	4	36.4	908	3	US-08-823-110-1	Sequence 1, Appl1	967	4	36.4	1346	2	US-08-470-566B-23	Sequence 23, Appl
895	4	36.4	908	3	US-08-604-296-1	Sequence 1, Appl1	968	4	36.4	1346	2	US-08-469-334-23	Sequence 23, Appl
896	4	36.4	908	4	US-08-855-146-2	Sequence 2, Appl1	969	4	36.4	1346	3	US-09-300-529-23	Sequence 23, Appl
897	4	36.4	909	5	US-08-310-912A-142	Sequence 142, App	970	4	36.4	1358	2	US-08-570-311-27	Sequence 27, Appl
898	4	36.4	909	5	PCT-US95-04589-142	Sequence 142, App	971	4	36.4	1439	2	US-08-449-644-4	Sequence 2, Appl1
899	4	36.4	913	3	US-08-911-853-17	Sequence 17, Appl	972	4	36.4	1439	2	US-08-087-244A-2	Sequence 2, Appl1
900	4	36.4	913	3	US-09-479-409-17	Sequence 17, Appl	973	4	36.4	1443	2	US-08-670-707A-39	Sequence 39, Appl
901	4	36.4	918	2	US-08-843-530B-35	Sequence 35, Appl	974	4	36.4	1443	2	US-09-037-601-39	Sequence 39, Appl
902	4	36.4	920	1	US-08-101-593-2	Sequence 2, Appl1	975	4	36.4	1457	2	US-08-652-971-3	Sequence 3, Appl1
903	4	36.4	920	1	US-08-101-593-4	Sequence 4, Appl1	976	4	36.4	1457	2	US-08-449-644-1	Sequence 1, Appl1



685	4	36.4	600	2	US-08-370-156-4	Sequence 4, Appl1	758	4	36.4	725	3	US-08-425-843-2	Sequence 2, Appl1
686	4	36.4	600	3	US-08-814-095-4	Sequence 2, Appl1	759	4	36.4	731	2	US-08-911-364-1	Sequence 1, Appl1
687	4	36.4	600	3	US-08-904-871-2	Sequence 4, Appl1	760	4	36.4	733	4	US-08-464-700-2	Sequence 2, Appl1
688	4	36.4	605	1	US-08-462-884A-3	Sequence 3, Appl1	761	4	734	4	US-09-000-016-4	Sequence 4, Appl1	
689	4	36.4	605	1	US-08-461-881B-3	Sequence 3, Appl1	762	4	36.4	737	1	US-08-185-432-2	Sequence 2, Appl1
690	4	36.4	605	2	US-08-889-402-2	Sequence 3, Appl1	763	4	36.4	737	1	US-08-185-432-4	Sequence 4, Appl1
691	4	36.4	605	2	US-08-687-956A-1	Sequence 1, Appl1	764	4	36.4	739	1	US-07-803-622E-2	Sequence 2, Appl1
692	4	36.4	605	2	US-09-123-960-3	Sequence 3, Appl1	765	4	36.4	739	2	US-08-836-943-2	Sequence 2, Appl1
693	4	36.4	607	2	US-08-752-307B-12	Sequence 12, Appl1	766	4	36.4	754	3	US-08-646-273-30	Sequence 30, Appl1
694	4	36.4	610	2	US-08-365-470-3	Sequence 3, Appl1	767	4	36.4	755	3	US-09-071-101-2	Sequence 2, Appl1
695	4	36.4	610	2	US-08-455-073A-2	Sequence 2, Appl1	768	4	36.4	755	3	US-09-369-618-2	Sequence 2, Appl1
696	4	36.4	610	3	US-09-209-668-19	Sequence 19, Appl1	769	4	36.4	755	3	US-09-369-617-2	Sequence 2, Appl1
697	4	36.4	610	6	5217870-2	Patent No. 5217870	770	4	36.4	758	1	US-08-289-112-2	Sequence 2, Appl1
698	4	36.4	614	1	US-07-732-962A-2	Sequence 2, Appl1	771	4	36.4	760	1	US-08-195-152-2	Sequence 2, Appl1
699	4	36.4	614	1	US-08-652-307A-2	Sequence 2, Appl1	772	4	36.4	761	4	US-09-046-894-35	Sequence 35, Appl1
700	4	36.4	614	2	US-08-370-156-2	Sequence 2, Appl1	773	4	36.4	763	2	US-08-677-862-2	Sequence 2, Appl1
701	4	36.4	614	3	US-08-446-100-19	Sequence 19, Appl1	774	4	36.4	763	2	US-09-252-571-2	Sequence 2, Appl1
702	4	36.4	614	3	US-08-446-100-20	Sequence 20, Appl1	775	4	36.4	763	3	US-09-434-065-2	Sequence 2, Appl1
703	4	36.4	614	3	US-08-446-100-21	Sequence 21, Appl1	776	4	36.4	763	4	US-08-789-275-4	Sequence 4, Appl1
704	4	36.4	614	3	US-08-446-100-22	Sequence 22, Appl1	777	4	36.4	763	4	US-08-789-275-5	Sequence 5, Appl1
705	4	36.4	614	3	US-08-446-100-23	Sequence 23, Appl1	778	4	36.4	768	2	US-08-232-617A-5	Sequence 5, Appl1
706	4	36.4	614	3	US-08-446-100-25	Sequence 25, Appl1	779	4	36.4	770	1	US-08-445-135-2	Sequence 2, Appl1
707	4	36.4	614	3	US-08-814-095-2	Sequence 25, Appl1	780	4	36.4	776	4	US-08-961-083-86	Sequence 86, Appl1
708	4	36.4	614	5	PCT-US92-06106-2	Sequence 2, Appl1	781	4	36.4	777	4	US-08-764-870-13	Sequence 13, Appl1
709	4	36.4	616	1	US-08-385-370-2	Sequence 2, Appl1	782	4	36.4	782	1	US-07-725-083-2	Sequence 2, Appl1
710	4	36.4	616	1	US-08-385-370-4	Sequence 4, Appl1	783	4	36.4	782	4	US-08-669-886-10	Sequence 10, Appl1
711	4	36.4	617	2	US-08-370-156-6	Sequence 6, Appl1	784	4	36.4	782	4	US-09-469-253-10	Sequence 10, Appl1
712	4	36.4	617	3	US-08-814-095-6	Sequence 6, Appl1	785	4	36.4	791	4	US-08-861-745B-1	Sequence 1, Appl1
713	4	36.4	622	1	US-08-426-819A-35	Sequence 35, Appl1	786	4	36.4	792	2	US-08-678-039A-40	Sequence 40, Appl1
714	4	36.4	624	2	US-08-879-561-3	Sequence 3, Appl1	787	4	36.4	792	2	US-08-190-802A-64	Sequence 64, Appl1
715	4	36.4	624	2	US-08-879-561-9	Sequence 9, Appl1	788	4	36.4	798	1	US-08-190-802A-68	Sequence 68, Appl1
716	4	36.4	628	3	US-08-776-271-2	Sequence 2, Appl1	789	4	36.4	798	2	US-08-308-818-2	Sequence 2, Appl1
717	4	36.4	628	4	US-09-215-035-2	Sequence 2, Appl1	790	4	36.4	805	4	US-08-989-299-2	Sequence 2, Appl1
718	4	36.4	629	3	US-08-556-419-23	Sequence 23, Appl1	791	4	36.4	811	2	US-08-933-750C-45	Sequence 45, Appl1
719	4	36.4	630	2	US-08-797-366-3	Sequence 3, Appl1	792	4	36.4	811	4	US-09-234-613-45	Sequence 45, Appl1
720	4	36.4	630	2	US-08-956-268-3	Sequence 3, Appl1	793	4	36.4	814	1	US-08-318-831-2	Sequence 2, Appl1
721	4	36.4	635	4	US-08-857-076-101	Sequence 101, Appl1	794	4	36.4	816	1	US-07-731-157A-4	Sequence 4, Appl1
722	4	36.4	638	2	US-08-426-125-4	Sequence 4, Appl1	795	4	36.4	816	1	US-08-229-444B-2	Sequence 2, Appl1
723	4	36.4	638	2	US-08-455-355-4	Sequence 4, Appl1	796	4	36.4	816	2	US-08-267-803B-9	Sequence 9, Appl1
724	4	36.4	639	1	US-08-466-390-2	Sequence 2, Appl1	797	4	36.4	816	2	US-08-541-180-4	Sequence 4, Appl1
725	4	36.4	639	1	US-08-470-850-2	Sequence 2, Appl1	798	4	36.4	816	4	US-09-041-886-17	Sequence 17, Appl1
726	4	36.4	639	1	US-08-467-781-2	Sequence 2, Appl1	799	4	36.4	820	4	US-09-173-914-2	Sequence 2, Appl1
727	4	36.4	639	1	US-08-195-487-2	Sequence 2, Appl1	800	4	36.4	821	1	US-07-935-311A-4	Sequence 4, Appl1
728	4	36.4	639	2	US-08-483-924-2	Sequence 2, Appl1	801	4	36.4	821	1	US-08-368-079-4	Sequence 4, Appl1
729	4	36.4	639	5	PCT-US93-06160-2	Sequence 2, Appl1	802	4	36.4	821	5	PCT-US93-07996-4	Sequence 2, Appl1
730	4	36.4	640	4	US-09-177-349-4	Sequence 4, Appl1	803	4	36.4	823	4	US-09-000-0163-2	Sequence 2, Appl1
731	4	36.4	642	2	US-08-245-511-48	Sequence 48, Appl1	804	4	36.4	828	3	US-08-481-435-2	Sequence 2, Appl1
732	4	36.4	642	2	US-08-600-993A-48	Sequence 48, Appl1	805	4	36.4	830	1	US-07-977-334-6	Sequence 6, Appl1
733	4	36.4	645	1	US-07-779-172A-3	Sequence 3, Appl1	806	4	36.4	830	1	US-08-458-819-6	Sequence 6, Appl1
734	4	36.4	648	4	US-09-183-706-43	Sequence 43, Appl1	807	4	36.4	830	5	PCT-US91-07035-6	Sequence 6, Appl1
735	4	36.4	649	4	US-09-188-930-305	Sequence 305, Appl1	808	4	36.4	834	1	US-08-471-033-21	Sequence 21, Appl1
736	4	36.4	655	1	US-07-736-178C-2	Sequence 2, Appl1	809	4	36.4	834	2	US-08-471-033-21	Sequence 21, Appl1
737	4	36.4	657	3	US-08-883-852A-3	Sequence 3, Appl1	810	4	36.4	834	2	US-08-463-884A-21	Sequence 21, Appl1
738	4	36.4	657	3	US-07-705-490-14	Sequence 14, Appl1	811	4	36.4	834	2	US-08-471-046A-21	Sequence 21, Appl1
739	4	36.4	657	4	US-08-821-818-3	Sequence 3, Appl1	812	4	36.4	834	2	US-08-470-566B-21	Sequence 21, Appl1
740	4	36.4	657	4	US-07-751-891B-14	Sequence 14, Appl1	813	4	36.4	834	2	US-08-469-334-21	Sequence 21, Appl1
741	4	36.4	664	1	US-08-421-661-6	Sequence 6, Appl1	814	4	36.4	837	3	US-09-300-529-21	Sequence 21, Appl1
742	4	36.4	664	3	US-08-669-408B-2	Sequence 2, Appl1	815	4	36.4	837	1	US-07-923-876-2	Sequence 2, Appl1
743	4	36.4	666	1	US-08-318-831-3	Sequence 3, Appl1	816	4	36.4	845	3	US-08-804-439A-94	Sequence 94, Appl1
744	4	36.4	668	1	US-08-205-018-2	Sequence 2, Appl1	817	4	36.4	845	3	US-08-720-229-94	Sequence 94, Appl1
745	4	36.4	669	2	US-07-861-800-2	Sequence 2, Appl1	818	4	36.4	846	3	US-08-885-291-55	Sequence 55, Appl1
746	4	36.4	694	2	US-08-895-522-4	Sequence 4, Appl1	819	4	36.4	846	3	US-09-107-847-2	Sequence 2, Appl1
747	4	36.4	694	2	US-09-195-391-4	Sequence 4, Appl1	820	4	36.4	849	1	US-08-405-392-18	Sequence 18, Appl1
748	4	36.4	694	3	US-08-559-397A-31	Sequence 31, Appl1	821	4	36.4	849	3	US-08-804-439A-17	Sequence 17, Appl1
749	4	36.4	699	1	US-08-348-006B-7	Sequence 7, Appl1	822	4	36.4	849	3	US-08-487-691-18	Sequence 18, Appl1
750	4	36.4	699	2	US-08-800-825A-7	Sequence 7, Appl1	823	4	36.4	849	3	US-08-720-229-17	Sequence 17, Appl1
751	4	36.4	699	4	US-09-158-657-7	Sequence 7, Appl1	824	4	36.4	850	3	US-08-481-735-10	Sequence 10, Appl1
752	4	36.4	699	4	PCT-US94-07297-39	Sequence 39, Appl1	825	4	36.4	851	2	US-08-369-996-2	Sequence 2, Appl1
753	4	36.4	700	3	US-08-931-952-2	Sequence 2, Appl1	826	4	36.4	851	2	US-08-852-091-2	Sequence 2, Appl1
754	4	36.4	700	3	US-08-272-247-2	Sequence 2, Appl1	827	4	36.4	851	3	US-08-830-754-2	Sequence 2, Appl1
755	4	36.4	700	5	PCT-US95-08560-2	Sequence 2, Appl1	828	4	36.4	851	3	US-08-956-652-2	Sequence 2, Appl1
756	4	36.4	708	3	US-08-646-273-23	Sequence 23, Appl1	829	4	36.4	851	3	US-08-956-669-9	Sequence 2, Appl1
757	4	36.4	715	3	US-08-425-843-7	Sequence 7, Appl1	830	4	36.4	851	3	US-09-012-710-2	Sequence 2, Appl1

539	4	36.4	428	1	US-08-332-576-2	Sequence 2, Appl1	612	4	36.4	500	2	US-08-987-519-1	Sequence 1, Appl1
540	4	36.4	428	1	US-08-551-687-1	Sequence 1, Appl1	613	4	36.4	501	2	PCT-US95-05471-77	Sequence 77, Appl
541	4	36.4	428	5	PCT-US95-13672-2	Sequence 2, Appl1	614	4	36.4	500	2	US-08-966-650-4	Sequence 4, Appl1
542	4	36.4	431	3	US-08-807-342B-5	Sequence 5, Appl1	615	4	36.4	503	3	US-09-010-998-2	Sequence 2, Appl1
543	4	36.4	431	4	US-09-188-930-3A1	Sequence 3A1, App	616	4	36.4	506	2	US-08-849-460A-5	Sequence 3, Appl1
544	4	36.4	431	4	US-09-381-681-3	Sequence 3, Appl1	617	4	36.4	507	1	US-08-305-505-3	Sequence 5, Appl1
545	4	36.4	432	1	US-08-522-166-8	Sequence 8, Appl1	618	4	36.4	507	6	US-08-849-460A-5	Sequence 3, Appl1
546	4	36.4	432	1	US-08-488-382A-8	Sequence 8, Appl1	619	4	36.4	509	1	US-08-030-096-2	Sequence 2, Appl1
547	4	36.4	432	2	US-08-480-912-8	Sequence 8, Appl1	620	4	36.4	509	2	US-08-559-505-2	Sequence 2, Appl1
548	4	36.4	432	5	PCT-US95-04910-13	Sequence 13, Appl1	621	4	36.4	509	2	US-08-749-907-2	Sequence 2, Appl1
549	4	36.4	433	1	US-07-672-483-4	Sequence 4, Appl1	622	4	36.4	509	2	US-08-890-960-4	Sequence 4, Appl1
550	4	36.4	433	3	US-08-941-533-6	Sequence 6, Appl1	623	4	36.4	509	3	US-08-890-979-4	Sequence 4, Appl1
551	4	36.4	434	3	US-08-337-602-3	Sequence 3, Appl1	624	4	36.4	509	3	US-08-822-324-8	Sequence 8, Appl1
552	4	36.4	434	3	US-08-558-135-3	Sequence 3, Appl1	625	4	36.4	509	4	US-09-033-894-4	Sequence 4, Appl1
553	4	36.4	437	1	US-08-136-119-2	Sequence 2, Appl1	626	4	36.4	509	4	US-08-857-076-46	Sequence 46, Appl1
554	4	36.4	437	2	US-08-481-814A-7	Sequence 7, Appl1	627	4	36.4	509	4	US-09-031-626-4	Sequence 4, Appl1
555	4	36.4	438	1	US-07-923-095-2	Sequence 2, Appl1	628	4	36.4	509	5	PCT-US95-07721-4	Sequence 4, Appl1
556	4	36.4	438	1	US-08-229-511-2	Sequence 2, Appl1	629	4	36.4	517	1	US-08-132-649-2	Sequence 2, Appl1
557	4	36.4	438	1	US-08-314-979-2	Sequence 2, Appl1	630	4	36.4	517	1	US-08-764-343-4	Sequence 4, Appl1
558	4	36.4	438	1	US-08-436-716-2	Sequence 2, Appl1	631	4	36.4	517	3	US-08-767-579-2	Sequence 2, Appl1
559	4	36.4	440	1	US-07-930-686-12	Sequence 12, Appl1	632	4	36.4	520	1	US-08-261-832A-10	Sequence 10, Appl1
560	4	36.4	440	2	US-08-460-998-12	Sequence 12, Appl1	633	4	36.4	520	1	US-09-000-016-7	Sequence 7, Appl1
561	4	36.4	441	4	US-08-244-603A-1	Sequence 1, Appl1	634	4	36.4	520	5	PCT-US95-07744A-10	Sequence 10, Appl1
562	4	36.4	441	4	US-09-191-136-31	Sequence 31, Appl1	635	4	36.4	525	4	US-08-984-618-17	Sequence 17, Appl1
563	4	36.4	443	1	US-08-476-008-69	Sequence 69, Appl1	636	4	36.4	529	1	US-07-779-890-2	Sequence 2, Appl1
564	4	36.4	443	1	US-08-306-063-69	Sequence 69, Appl1	637	4	36.4	529	1	US-07-779-890-2	Sequence 2, Appl1
565	4	36.4	443	1	US-08-833-485-69	Sequence 69, Appl1	638	4	36.4	529	5	PCT-US93-05640-2	Sequence 2, Appl1
566	4	36.4	443	1	US-09-137-440-69	Sequence 69, Appl1	639	4	36.4	537	1	US-08-173-508-2	Sequence 2, Appl1
567	4	36.4	445	1	US-08-447-925-6	Sequence 6, Appl1	640	4	36.4	537	1	US-08-472-028A-2	Sequence 2, Appl1
568	4	36.4	445	1	US-08-353-400-33	Sequence 33, Appl1	641	4	36.4	537	2	US-08-265-310-2	Sequence 2, Appl1
569	4	36.4	447	6	US-09-109-204-3	Sequence 3, Appl1	642	4	36.4	537	2	US-08-808-931-2	Sequence 2, Appl1
570	4	36.4	447	6	5455030-1	Sequence 21, Appl1	643	4	36.4	537	3	US-08-808-931-2	Sequence 2, Appl1
571	4	36.4	448	4	US-08-952-127-21	Sequence 38, Appl1	644	4	36.4	537	3	US-09-050-603A-2	Sequence 2, Appl1
572	4	36.4	453	4	US-08-961-083-38	Sequence 10, Appl1	645	4	36.4	537	3	US-09-102-420B-2	Sequence 2, Appl1
573	4	36.4	454	1	US-07-930-686-10	Sequence 10, Appl1	646	4	36.4	537	3	US-08-951-742-2	Sequence 2, Appl1
574	4	36.4	454	2	US-08-460-998-10	Sequence 10, Appl1	647	4	36.4	537	4	US-09-071-296-2	Sequence 2, Appl1
575	4	36.4	455	2	US-08-272-255-14	Sequence 14, Appl1	648	4	36.4	538	2	US-08-729-214-24	Sequence 24, Appl1
576	4	36.4	455	5	PCT-US95-08565-14	Sequence 14, Appl1	649	4	36.4	538	3	US-09-028-994-24	Sequence 24, Appl1
577	4	36.4	455	1	US-08-403-545-5	Sequence 5, Appl1	650	4	36.4	539	3	US-09-264-737-1	Sequence 1, Appl1
578	4	36.4	459	4	US-08-404-381-5	Sequence 5, Appl1	651	4	36.4	540	4	US-08-991-677-8	Sequence 8, Appl1
579	4	36.4	461	1	US-08-672-571A-3	Sequence 3, Appl1	652	4	36.4	545	1	US-08-133-347-2	Sequence 2, Appl1
580	4	36.4	464	1	US-08-353-400-36	Sequence 36, Appl1	653	4	36.4	548	4	US-08-687-590-31	Sequence 31, Appl1
581	4	36.4	469	1	US-08-551-687-6	Sequence 6, Appl1	654	4	36.4	549	5	PCT-US93-05701-2	Sequence 2, Appl1
582	4	36.4	469	2	US-08-551-687-6	Sequence 6, Appl1	655	4	36.4	550	3	US-08-807-342B-4	Sequence 4, Appl1
583	4	36.4	471	3	US-08-911-853-4	Sequence 4, Appl1	656	4	36.4	554	4	US-08-180-371-2	Sequence 2, Appl1
584	4	36.4	471	4	US-09-479-409-4	Sequence 4, Appl1	657	4	36.4	557	4	US-09-027-064-2	Sequence 2, Appl1
585	4	36.4	472	4	US-09-088-425-1	Sequence 1, Appl1	658	4	36.4	559	1	US-08-424-788-3	Sequence 3, Appl1
586	4	36.4	478	5	PCT-US95-08493-15	Sequence 15, Appl1	659	4	36.4	559	2	US-08-756-317-10	Sequence 10, Appl1
587	4	36.4	480	2	US-08-570-227A-2	Sequence 2, Appl1	660	4	36.4	562	2	US-08-687-702-1	Sequence 1, Appl1
588	4	36.4	480	4	US-09-077-991-2	Sequence 2, Appl1	661	4	36.4	566	3	US-09-264-737-2	Sequence 2, Appl1
589	4	36.4	485	1	US-07-881-075-1	Sequence 1, Appl1	662	4	36.4	567	3	US-08-646-273-19	Sequence 19, Appl1
590	4	36.4	485	1	US-08-120-827-1	Sequence 1, Appl1	663	4	36.4	568	1	US-07-803-632E-4	Sequence 4, Appl1
591	4	36.4	485	1	US-08-453-956-15	Sequence 15, Appl1	664	4	36.4	568	1	US-08-262-424-7	Sequence 7, Appl1
592	4	36.4	485	1	US-08-478-675-1	Sequence 1, Appl1	665	4	36.4	568	2	US-08-493-197-7	Sequence 7, Appl1
593	4	36.4	485	1	US-08-086-631-15	Sequence 15, Appl1	666	4	36.4	568	2	US-08-717-587A-4	Sequence 4, Appl1
594	4	36.4	485	2	US-08-452-930-15	Sequence 15, Appl1	667	4	36.4	568	3	US-08-883-610A-4	Sequence 4, Appl1
595	4	36.4	485	5	PCT-US93-08174-15	Sequence 15, Appl1	668	4	36.4	568	4	US-08-936-094A-4	Sequence 4, Appl1
596	4	36.4	489	1	US-08-318-831-4	Sequence 4, Appl1	669	4	36.4	568	5	PCT-US95-07844-7	Sequence 7, Appl1
597	4	36.4	489	1	PCT-US93-11110-1	Sequence 1, Appl1	670	4	36.4	575	1	US-08-424-788-2	Sequence 2, Appl1
598	4	36.4	489	6	5221789-1	Sequence 1, Appl1	671	4	36.4	575	1	US-08-110-663-4	Sequence 4, Appl1
599	4	36.4	490	2	US-08-687-916-24	Sequence 24, Appl1	672	4	36.4	575	2	US-08-477-166-4	Sequence 4, Appl1
600	4	36.4	490	2	US-09-138-614-24	Sequence 24, Appl1	673	4	36.4	575	2	US-08-472-097-4	Sequence 4, Appl1
601	4	36.4	490	4	US-08-687-916-23	Sequence 23, Appl1	674	4	36.4	575	5	PCT-US93-11638-4	Sequence 4, Appl1
602	4	36.4	491	2	US-08-687-916-23	Sequence 23, Appl1	675	4	36.4	584	1	US-08-426-819A-36	Sequence 36, Appl1
603	4	36.4	491	4	US-09-138-614-23	Sequence 23, Appl1	676	4	36.4	590	3	US-08-893-852A-4	Sequence 4, Appl1
604	4	36.4	493	4	US-08-378-313-19	Sequence 19, Appl1	677	4	36.4	590	4	US-08-821-818-2	Sequence 2, Appl1
605	4	36.4	493	4	US-08-378-313-21	Sequence 21, Appl1	678	4	36.4	591	2	US-08-889-402-1	Sequence 1, Appl1
606	4	36.4	493	4	US-08-378-313-28	Sequence 28, Appl1	679	4	36.4	594	3	US-08-581-148C-14	Sequence 14, Appl1
607	4	36.4	494	4	US-08-378-313-23	Sequence 23, Appl1	680	4	36.4	597	1	US-08-462-884A-1	Sequence 1, Appl1
608	4	36.4	494	4	US-08-378-313-23	Sequence 23, Appl1	681	4	36.4	597	1	US-08-461-881B-1	Sequence 1, Appl1
609	4	36.4	495	2	US-08-687-916-22	Sequence 22, Appl1	682	4	36.4	597	2	US-09-123-960-1	Sequence 1, Appl1
610	4	36.4	495	4	US-09-138-614-22	Sequence 22, Appl1	683	4	36.4	599	1	US-08-442-542-18	Sequence 18, Appl1
611	4	36.4	500	1	US-08-260-582-77	Sequence 77, Appl1	684	4	36.4	599	3	US-08-765-469-18	Sequence 18, Appl1

393	4	36.4	296	1	US-07-745-382-14	Sequence 14, Appl	466	4	36.4	362	2	US-09-076-193-3	Sequence 3, Appl1
394	4	36.4	296	1	US-07-921-848-14	Sequence 14, Appl	467	4	36.4	362	2	US-09-076-193-5	Sequence 7, Appl1
395	4	36.4	296	1	US-08-115-680-4	Sequence 4, Appl1	468	4	36.4	362	2	US-09-076-193-7	Sequence 11, Appl
396	4	36.4	296	1	US-07-941-372-4	Sequence 14, Appl	469	4	36.4	362	5	PCT-US94-05288-11	Sequence 2, Appl1
397	4	36.4	296	1	US-08-165-301A-14	Sequence 14, Appl	470	4	36.4	363	3	US-08-978-589A-2	Sequence 1, Appl1
398	4	36.4	296	4	US-08-810-436-14	Sequence 14, Appl	471	4	36.4	363	4	US-09-336-601-1	Sequence 1, Appl1
399	4	36.4	296	4	PCT-US93-08247-4	Sequence 4, Appl1	472	4	36.4	363	4	US-09-301-665-4	Sequence 4, Appl1
400	4	36.4	296	5	PCT-US94-14179-14	Sequence 14, Appl	473	4	36.4	364	1	US-08-626-994A-1	Sequence 1, Appl1
401	4	36.4	297	4	US-08-580-545B-6	Sequence 6, Appl1	474	4	36.4	364	4	US-08-957-442-1	Sequence 1, Appl1
402	4	36.4	297	4	US-09-262-653A-6	Sequence 6, Appl1	475	4	36.4	364	5	PCT-US96-10618-2	Sequence 2, Appl1
403	4	36.4	298	4	US-08-961-083-24	Sequence 24, Appl	476	4	36.4	365	2	US-08-515-251A-2	Sequence 2, Appl1
404	4	36.4	299	2	US-08-773-608A-2	Sequence 20, Appl	477	4	36.4	366	3	US-08-945-056-6	Sequence 8, Appl1
405	4	36.4	302	3	US-08-303-861-20	Sequence 2, Appl1	478	4	36.4	367	2	US-08-945-056-8	Sequence 4, Appl1
406	4	36.4	308	1	US-08-499-568-2	Sequence 2, Appl1	479	4	36.4	367	2	US-08-515-251A-4	Sequence 1, Appl1
407	4	36.4	308	2	US-08-793-958-2	Sequence 2, Appl1	480	4	36.4	369	2	US-08-139-609-1	Sequence 1, Appl1
408	4	36.4	309	4	US-09-342-084-2	Sequence 2, Appl1	481	4	36.4	375	1	US-08-027-986-1	Sequence 1, Appl1
409	4	36.4	315	5	PCT-US94-03796-2	Sequence 2, Appl1	482	4	36.4	375	1	US-08-027-986-2	Sequence 2, Appl1
410	4	36.4	317	1	US-07-688-299-1	Sequence 1, Appl1	483	4	36.4	375	2	US-08-525-596B-14	Sequence 5, Appl1
411	4	36.4	317	1	US-07-980-517A-1	Sequence 1, Appl1	484	4	36.4	375	2	US-08-765-875-5	Sequence 5, Appl1
412	4	36.4	318	1	US-07-688-299-3	Sequence 3, Appl1	485	4	36.4	375	3	US-08-795-671-5	Sequence 5, Appl1
413	4	36.4	318	1	US-07-688-299-13	Sequence 13, Appl	486	4	36.4	375	3	US-09-177-860A-14	Sequence 14, Appl
414	4	36.4	319	2	US-08-347-335A-3	Sequence 3, Appl	487	4	36.4	377	1	US-08-891-789B-2	Sequence 2, Appl1
415	4	36.4	322	4	US-09-188-930-141	Sequence 141, App	488	4	36.4	377	1	US-08-525-697-2	Sequence 2, Appl1
416	4	36.4	324	2	US-08-579-940-7	Sequence 7, Appl1	489	4	36.4	379	4	US-09-191-136-32	Sequence 32, Appl
417	4	36.4	325	2	US-09-018-576-3	Sequence 3, Appl1	490	4	36.4	379	4	US-09-426-557-2	Sequence 2, Appl1
418	4	36.4	325	2	US-09-018-576-12	Sequence 12, Appl	491	4	36.4	379	4	US-09-426-557-4	Sequence 4, Appl1
419	4	36.4	325	3	US-09-248-137-3	Sequence 3, Appl	492	4	36.4	379	4	US-09-426-557-6	Sequence 6, Appl1
420	4	36.4	325	3	US-09-248-137-12	Sequence 12, Appl	493	4	36.4	379	4	US-09-426-557-8	Sequence 8, Appl1
421	4	36.4	330	2	US-08-838-543-3	Sequence 3, Appl1	494	4	36.4	380	4	US-08-924-545-2	Sequence 4, Appl1
422	4	36.4	331	3	US-08-445-515-59	Sequence 59, Appl	495	4	36.4	382	2	US-08-455-968E-5	Sequence 5, Appl1
423	4	36.4	335	4	US-09-387-699-2	Sequence 2, Appl1	496	4	36.4	382	2	US-08-360-606B-30	Sequence 30, Appl
424	4	36.4	336	4	US-08-235-836C-54	Sequence 54, Appl	497	4	36.4	382	2	US-08-823-116-139	Sequence 139, App
425	4	36.4	338	1	US-08-602-359A-40	Sequence 40, Appl	498	4	36.4	383	6	5470718-5	Patent No. 5470718
426	4	36.4	339	3	US-08-626-994A-3	Sequence 3, Appl1	499	4	36.4	384	2	US-08-637-559B-375	Sequence 375, App
427	4	36.4	339	3	US-08-957-742-3	Sequence 3, Appl1	500	4	36.4	384	3	US-08-871-335A-375	Sequence 375, App
428	4	36.4	339	4	US-09-330-117B-4	Sequence 10, Appl	501	4	36.4	388	1	US-08-499-568-11	Sequence 11, Appl
429	4	36.4	339	4	US-09-330-117B-10	Sequence 10, Appl	502	4	36.4	388	2	US-08-793-958-11	Sequence 11, Appl
430	4	36.4	339	4	US-09-330-117B-12	Sequence 12, Appl	503	4	36.4	391	2	US-08-244-951A-10	Sequence 10, Appl
431	4	36.4	339	4	US-09-330-117B-14	Sequence 14, Appl	504	4	36.4	391	2	US-08-389-011-123	Sequence 23, Appl
432	4	36.4	339	4	US-09-330-117B-16	Sequence 16, Appl	505	4	36.4	391	4	US-08-403-017A-23	Sequence 23, Appl
433	4	36.4	340	2	US-08-446-875-15	Sequence 15, Appl	506	4	36.4	391	4	US-09-348-952A-23	Sequence 23, Appl
434	4	36.4	340	4	US-08-960-780-34	Sequence 34, Appl	507	4	36.4	393	6	5182195-10	Patent No. 5182195
435	4	36.4	340	4	US-09-073-898-34	Sequence 34, Appl	508	4	36.4	394	4	US-08-499-568-4	Sequence 4, Appl1
436	4	36.4	344	1	US-08-180-209B-58	Sequence 58, Appl	509	4	36.4	394	2	US-08-793-958-4	Sequence 4, Appl1
437	4	36.4	344	2	US-08-755-728-3	Sequence 3, Appl1	510	4	36.4	406	1	US-07-973-431B-1	Sequence 1, Appl1
438	4	36.4	344	2	US-08-974-655-3	Sequence 3, Appl1	511	4	36.4	407	2	US-08-765-875-2	Sequence 2, Appl1
439	4	36.4	344	5	US-09-283-011-1-3	Sequence 3, Appl1	512	4	36.4	407	2	US-08-765-875-6	Sequence 6, Appl1
440	4	36.4	346	2	PCT-US94-02629-58	Sequence 58, Appl	513	4	36.4	407	2	US-08-776-885-3	Sequence 3, Appl1
441	4	36.4	346	2	US-08-401-068-8	Sequence 8, Appl1	514	4	36.4	407	3	US-08-795-671-2	Sequence 2, Appl1
442	4	36.4	346	2	US-08-846-338-8	Sequence 36, Appl	515	4	36.4	408	2	US-08-683-007A-2	Sequence 2, Appl1
443	4	36.4	346	2	US-08-687-702-36	Sequence 2, Appl1	516	4	36.4	408	2	US-08-809-440A-5	Sequence 5, Appl1
444	4	36.4	346	3	US-08-411-768B-2	Sequence 2, Appl1	517	4	36.4	412	2	US-08-463-081B-14	Sequence 14, Appl
445	4	36.4	347	2	US-09-016-000-1	Sequence 2, Appl1	518	4	36.4	412	2	US-08-461-979A-14	Sequence 14, Appl
446	4	36.4	347	4	US-08-960-780-25	Sequence 25, Appl	519	4	36.4	412	2	US-08-462-390B-14	Sequence 14, Appl
447	4	36.4	347	4	US-09-073-898-25	Sequence 25, Appl	520	4	36.4	412	2	US-08-463-074B-14	Sequence 14, Appl
448	4	36.4	348	4	US-08-960-780-21	Sequence 21, Appl	521	4	36.4	412	3	US-08-465-855C-14	Sequence 14, Appl
449	4	36.4	348	4	US-08-960-780-42	Sequence 42, Appl	522	4	36.4	412	3	US-08-652-446-14	Sequence 14, Appl
450	4	36.4	348	4	US-09-073-898-21	Sequence 21, Appl	523	4	36.4	412	4	US-09-027-064-4	Sequence 4, Appl1
451	4	36.4	351	1	US-08-159-969-2	Sequence 42, Appl	524	4	36.4	412	4	US-08-669-108B-10	Sequence 10, Appl
452	4	36.4	351	2	US-08-726-306A-17	Sequence 17, Appl	525	4	36.4	415	2	US-08-576-626A-52	Sequence 52, Appl
453	4	36.4	352	2	US-08-137-627-4	Sequence 4, Appl1	526	4	36.4	417	4	US-09-276-400-5	Sequence 5, Appl1
454	4	36.4	359	2	US-08-865-348-4	Sequence 4, Appl1	527	4	36.4	419	3	US-09-155-200-4	Sequence 4, Appl1
455	4	36.4	359	2	US-08-809-103B-2	Sequence 2, Appl1	528	4	36.4	422	1	US-08-132-649-4	Sequence 4, Appl1
456	4	36.4	359	4	US-08-809-103B-4	Sequence 4, Appl1	529	4	36.4	422	1	US-08-167-579-4	Sequence 4, Appl1
457	4	36.4	359	4	US-08-809-103B-6	Sequence 6, Appl1	530	4	36.4	424	1	US-08-419-614-2	Sequence 2, Appl1
458	4	36.4	359	4	US-08-809-103B-8	Sequence 8, Appl1	531	4	36.4	424	2	US-08-871-268A-23	Sequence 23, Appl
459	4	36.4	361	2	US-08-209-521-5	Sequence 5, Appl1	532	4	36.4	424	3	US-08-871-267B-31	Sequence 31, Appl
460	4	36.4	361	2	US-09-034-985-2	Sequence 2, Appl1	533	4	36.4	425	4	US-08-960-780-27	Sequence 27, Appl
461	4	36.4	361	3	US-08-961-810-1	Sequence 1, Appl1	534	4	36.4	425	4	US-09-073-898-27	Sequence 27, Appl
462	4	36.4	361	4	US-08-352-902D-1	Sequence 11, Appl1	535	4	36.4	427	4	US-08-969-125-9	Sequence 9, Appl1
463	4	36.4	362	1	US-08-247-907A-11	Sequence 11, Appl	536	4	36.4	428	1	US-08-190-802A-29	Sequence 29, Appl
464	4	36.4	362	1	US-08-452-772-11	Sequence 11, Appl	537	4	36.4	428	1	US-08-353-550-1	Sequence 1, Appl1
465	4	36.4	362	1			538	4	36.4				

247	4	36.4	111	1	US-08-466-886-33	Sequence 33, Appl	320	4	36.4	188	2	US-08-933-750C-8	Sequence 8, Appl
248	4	36.4	111	2	US-08-852-091-18	Sequence 18, Appl	321	4	36.4	188	4	US-09-234-613-8	Sequence 8, Appl
249	4	36.4	111	2	US-08-665-202-36	Sequence 36, Appl	322	4	36.4	189	2	US-08-861-269-7	Sequence 7, Appl
250	4	36.4	111	2	US-08-665-202-40	Sequence 40, Appl	323	4	36.4	189	2	US-09-134-556-7	Sequence 7, Appl
251	4	36.4	111	2	US-08-665-202-41	Sequence 41, Appl	324	4	36.4	189	3	US-09-293-273-7	Sequence 7, Appl
252	4	36.4	111	2	US-08-665-202-42	Sequence 42, Appl	325	4	36.4	189	3	US-08-646-273-14	Sequence 14, Appl
253	4	36.4	111	2	US-08-665-202-43	Sequence 43, Appl	326	4	36.4	193	3	US-09-041-889-5	Sequence 5, Appl
254	4	36.4	111	4	US-08-983-607-21	Sequence 21, Appl	327	4	36.4	193	3	US-08-837-058-5	Sequence 5, Appl
255	4	36.4	111	4	US-08-983-607-23	Sequence 23, Appl	328	4	36.4	196	2	US-08-829-418-5	Sequence 5, Appl
256	4	36.4	111	4	US-08-983-607-35	Sequence 35, Appl	329	4	36.4	196	2	US-08-748-483-3	Sequence 3, Appl
257	4	36.4	111	4	US-08-469-617-33	Sequence 33, Appl	330	4	36.4	199	3	US-08-737-248-9	Sequence 9, Appl
258	4	36.4	111	5	PCT-US95-08950-6	Sequence 6, Appl	331	4	36.4	201	1	US-08-292-945-4	Sequence 4, Appl
259	4	36.4	111	5	PCT-US95-17025-18	Sequence 18, Appl	332	4	36.4	201	1	US-08-252-073A-4	Sequence 4, Appl
260	4	36.4	112	3	US-08-665-202-39	Sequence 39, Appl	333	4	36.4	201	5	PCT-US93-12074-4	Sequence 4, Appl
261	4	36.4	112	3	US-08-665-360-1	Sequence 1, Appl	334	4	36.4	205	2	US-08-912-227-4	Sequence 4, Appl
262	4	36.4	112	4	US-08-983-607-31	Sequence 31, Appl	335	4	36.4	206	4	US-09-107-383-15	Sequence 15, Appl
263	4	36.4	112	4	US-09-240-274-64	Sequence 64, Appl	336	4	36.4	209	4	US-09-266-220-1	Sequence 1, Appl
264	4	36.4	113	1	US-08-211-202-112	Sequence 112, App	337	4	36.4	209	4	US-09-266-220-2	Sequence 2, Appl
265	4	36.4	113	1	US-08-211-202-113	Sequence 113, App	338	4	36.4	214	2	US-08-572-447C-13	Sequence 13, Appl
266	4	36.4	113	1	US-08-211-202-114	Sequence 114, App	339	4	36.4	216	2	US-08-929-418-2	Sequence 2, Appl
267	4	36.4	114	2	US-09-240-274-62	Sequence 62, Appl	340	4	36.4	222	2	US-08-391-743A-2	Sequence 2, Appl
268	4	36.4	118	2	US-08-443-639-13	Sequence 13, Appl	341	4	36.4	226	2	US-08-572-447C-15	Sequence 15, Appl
269	4	36.4	118	3	US-08-577-483-13	Sequence 13, Appl	342	4	36.4	227	4	US-08-944-453-40	Sequence 10, Appl
270	4	36.4	119	2	US-08-588-258B-31	Sequence 31, Appl	343	4	36.4	233	2	US-09-151-611-1	Sequence 1, Appl
271	4	36.4	119	2	US-08-460-505-31	Sequence 31, Appl	344	4	36.4	233	4	US-08-523-854-6	Sequence 6, Appl
272	4	36.4	119	5	PCT-US96-08295-31	Sequence 31, Appl	345	4	36.4	234	3	US-08-487-550-2	Sequence 2, Appl
273	4	36.4	120	1	US-08-211-747-7	Sequence 7, Appl	346	4	36.4	235	2	US-08-378-939-12	Sequence 12, Appl
274	4	36.4	123	1	US-07-869-912-2	Sequence 2, Appl	347	4	36.4	235	4	US-09-049-672A-10	Sequence 10, Appl
275	4	36.4	123	1	US-08-482-728A-7	Sequence 7, Appl	348	4	36.4	235	4	US-09-049-672A-12	Sequence 12, Appl
276	4	36.4	126	1	US-08-202-047-13	Sequence 13, Appl	349	4	36.4	236	3	US-08-487-550-10	Sequence 10, Appl
277	4	36.4	126	1	US-08-202-047-15	Sequence 15, Appl	350	4	36.4	236	4	US-09-049-672A-8	Sequence 8, Appl
278	4	36.4	126	1	US-08-202-047-17	Sequence 17, Appl	351	4	36.4	236	5	PCT-US95-04801-8	Sequence 8, Appl
279	4	36.4	126	3	US-08-202-047-19	Sequence 19, Appl	352	4	36.4	237	2	US-08-464-517-24	Sequence 24, Appl
280	4	36.4	126	3	US-08-964-690-13	Sequence 13, Appl	353	4	36.4	237	2	US-08-246-361A-24	Sequence 24, Appl
281	4	36.4	126	3	US-08-964-690-15	Sequence 15, Appl	354	4	36.4	237	3	US-08-463-772-24	Sequence 24, Appl
282	4	36.4	126	3	US-08-964-690-17	Sequence 17, Appl	355	4	36.4	237	5	PCT-US93-05000-24	Sequence 24, Appl
283	4	36.4	126	3	US-08-964-690-19	Sequence 19, Appl	356	4	36.4	239	2	US-08-114-555A-12	Sequence 12, Appl
284	4	36.4	131	1	US-08-012-543-4	Sequence 4, Appl	357	4	36.4	239	3	US-08-559-397A-17	Sequence 17, Appl
285	4	36.4	131	1	US-08-305-683A-4	Sequence 4, Appl	358	4	36.4	240	4	US-09-049-672A-11	Sequence 11, Appl
286	4	36.4	131	1	US-08-371-131-25	Sequence 25, Appl	359	4	36.4	244	1	US-08-289-659A-3	Sequence 3, Appl
287	4	36.4	131	5	PCT-US93-07645A-4	Sequence 4, Appl	360	4	36.4	244	2	US-08-878-283-3	Sequence 3, Appl
288	4	36.4	132	2	US-08-345-321-4	Sequence 4, Appl	361	4	36.4	247	2	US-08-851-974-1	Sequence 1, Appl
289	4	36.4	132	2	US-08-649-991-95	Sequence 95, Appl	362	4	36.4	247	2	US-08-851-974-4	Sequence 4, Appl
290	4	36.4	133	1	US-08-196-218-32	Sequence 32, Appl	363	4	36.4	247	2	US-09-213-390-1	Sequence 1, Appl
291	4	36.4	133	1	US-08-681-953-32	Sequence 32, Appl	364	4	36.4	247	2	US-09-213-390-4	Sequence 4, Appl
292	4	36.4	133	2	US-08-649-991-110	Sequence 110, App	365	4	36.4	249	2	US-08-685-992-28	Sequence 28, Appl
293	4	36.4	134	2	US-08-850-910A-45	Sequence 45, Appl	366	4	36.4	249	2	US-09-144-925-28	Sequence 28, Appl
294	4	36.4	134	2	US-08-850-910A-48	Sequence 48, Appl	367	4	36.4	258	2	US-08-602-941-3	Sequence 3, Appl
295	4	36.4	138	2	US-08-684-101-2	Sequence 2, Appl	368	4	36.4	258	2	US-08-665-202-5	Sequence 5, Appl
296	4	36.4	138	2	US-09-205-814-2	Sequence 2, Appl	369	4	36.4	258	3	US-08-961-264-3	Sequence 3, Appl
297	4	36.4	147	3	US-09-136-442-3	Sequence 3, Appl	370	4	36.4	258	3	US-08-303-861-18	Sequence 18, Appl
298	4	36.4	148	5	PCT-US95-07135-2	Sequence 2, Appl	371	4	36.4	258	3	US-08-303-861-19	Sequence 19, Appl
299	4	36.4	152	1	US-08-046-583-11	Sequence 11, Appl	372	4	36.4	260	3	US-08-857-213-3	Sequence 3, Appl
300	4	36.4	152	1	US-08-261-677-11	Sequence 11, Appl	373	4	36.4	260	6	5320958-4	Patent No. 5320958
301	4	36.4	152	1	US-08-384-556A-2	Sequence 2, Appl	374	4	36.4	263	2	US-08-767-026-16	Sequence 13, Appl
302	4	36.4	152	1	US-08-384-556A-7	Sequence 7, Appl	375	4	36.4	264	1	US-08-463-115-93	Sequence 93, Appl
303	4	36.4	152	1	US-08-592-936B-17	Sequence 17, Appl	376	4	36.4	264	1	US-08-465-388-93	Sequence 93, Appl
304	4	36.4	152	1	US-08-331-355A-11	Sequence 11, Appl	377	4	36.4	268	1	US-08-446-919A-2	Sequence 2, Appl
305	4	36.4	152	2	US-08-331-355A-13	Sequence 23, Appl	378	4	36.4	270	4	US-09-082-953-10	Sequence 10, Appl
306	4	36.4	152	2	US-09-111-573-17	Sequence 17, Appl	379	4	36.4	270	4	US-09-342-084-8	Sequence 8, Appl
307	4	36.4	152	5	PCT-US94-12364-11	Sequence 11, Appl	380	4	36.4	277	6	5164481-2	Patent No. 5164481
308	4	36.4	152	5	PCT-US94-12364-23	Sequence 23, Appl	381	4	36.4	278	2	US-08-416-603-6	Sequence 6, Appl
309	4	36.4	152	5	PCT-US95-07753-2	Sequence 2, Appl	382	4	36.4	280	1	US-08-409-452-1	Sequence 1, Appl
310	4	36.4	152	5	PCT-US95-07753-7	Sequence 7, Appl	383	4	36.4	280	6	5182210-12	Patent No. 5182210
311	4	36.4	154	4	US-09-404-670-4	Sequence 4, Appl	384	4	36.4	281	3	US-09-320-878-5	Sequence 5, Appl
312	4	36.4	155	4	US-09-315-794-12	Sequence 12, Appl	385	4	36.4	286	3	US-08-891-789B-4	Sequence 4, Appl
313	4	36.4	155	4	US-09-389-341-12	Sequence 12, Appl	386	4	36.4	288	1	US-08-396-650-1	Sequence 1, Appl
314	4	36.4	155	4	US-09-389-341-12	Sequence 12, Appl	387	4	36.4	288	1	US-08-768-656-1	Sequence 1, Appl
315	4	36.4	163	3	US-09-053-197A-26	Sequence 26, Appl	388	4	36.4	291	4	US-09-328-869-2	Sequence 2, Appl
316	4	36.4	178	1	US-08-673-269-2	Sequence 2, Appl	389	4	36.4	293	4	US-09-037-143-52	Sequence 52, Appl
317	4	36.4	180	5	PCT-US93-04365-10	Sequence 10, Appl	390	4	36.4	293	4	US-08-260-174-52	Sequence 52, Appl
318	4	36.4	180	5	PCT-US93-04365-11	Sequence 11, Appl	391	4	36.4	294	2	US-08-923-856-3	Sequence 3, Appl
319	4	36.4	181	2	US-08-766-551-1	Sequence 1, Appl	392	4	36.4	294	3	US-09-216-294-3	Sequence 3, Appl

101	4	36.4	12	4	US-09-073-898-61	Sequence 61, Appl	174	4	36.4	69	1	US-08-175-393A-4	Sequence 4, Appl1
102	4	36.4	13	1	US-08-211-747-9	Sequence 29, Appl1	175	4	36.4	74	4	US-09-025-151-19	Sequence 12, Appl
103	4	36.4	13	2	US-08-666-473-29	Sequence 9, Appl	176	4	36.4	75	4	US-08-686-878A-12	Sequence 19, Appl
104	4	36.4	13	2	US-08-934-222-104	Sequence 104, App	177	4	36.4	79	1	US-08-011-398B-19	Sequence 19, Appl
105	4	36.4	13	2	US-08-933-402-104	Sequence 104, App	178	4	36.4	79	1	US-08-464-051-19	Sequence 19, Appl
106	4	36.4	13	2	US-09-207-621-104	Sequence 104, App	179	4	36.4	79	2	US-08-462-498-19	Sequence 19, Appl
107	4	36.4	13	2	US-08-532-818-104	Sequence 104, App	180	4	36.4	79	3	US-08-554-385-18	Sequence 18, Appl
108	4	36.4	13	3	US-09-231-797-104	Sequence 104, App	181	4	36.4	81	2	US-08-436-420-35	Sequence 35, Appl
109	4	36.4	13	3	US-08-934-224-104	Sequence 104, App	182	4	36.4	83	2	US-08-878-546-8	Sequence 8, Appl1
110	4	36.4	13	3	US-08-933-843-104	Sequence 104, App	183	4	36.4	98	2	US-08-211-202-111	Sequence 11, App
111	4	36.4	13	4	US-08-934-223-104	Sequence 104, App	184	4	36.4	98	2	US-08-665-202-37	Sequence 37, Appl
112	4	36.4	13	6	546585-11	Patent No. 546585	185	4	36.4	98	2	US-08-665-202-38	Sequence 38, Appl
113	4	36.4	14	1	US-07-841-997A-29	Sequence 29, Appl	186	4	36.4	99	3	US-09-173-581-8	Sequence 8, Appl1
114	4	36.4	14	1	US-08-290-301-29	Sequence 29, Appl	187	4	36.4	100	3	US-08-405-392-12	Sequence 12, Appl
115	4	36.4	15	2	US-08-850-910A-29	Sequence 29, Appl	188	4	36.4	100	3	US-08-487-691-12	Sequence 12, Appl
116	4	36.4	15	2	US-08-437-116A-5	Sequence 5, Appl1	189	4	36.4	103	1	US-07-942-245-8	Sequence 8, Appl
117	4	36.4	15	2	US-08-530-792D-5	Sequence 5, Appl1	190	4	36.4	103	1	US-08-107-669D-3	Sequence 3, Appl1
118	4	36.4	15	4	US-08-928-213B-18	Sequence 18, Appl	191	4	36.4	103	1	US-08-472-788A-3	Sequence 3, Appl1
119	4	36.4	16	1	US-07-942-245-95	Sequence 95, Appl	192	4	36.4	103	2	US-08-477-531B-3	Sequence 3, Appl1
120	4	36.4	16	1	US-07-942-245-99	Sequence 99, Appl	193	4	36.4	103	2	US-08-082-842A-3	Sequence 3, Appl1
121	4	36.4	16	1	US-07-942-245-100	Sequence 100, App	194	4	36.4	103	3	US-08-992-176-3	Sequence 3, Appl1
122	4	36.4	16	1	US-07-942-245-101	Sequence 101, App	195	4	36.4	103	4	US-09-413-814-38	Sequence 38, Appl
123	4	36.4	16	1	US-07-942-245-102	Sequence 102, App	196	4	36.4	103	4	US-09-240-274-42	Sequence 42, Appl
124	4	36.4	16	1	US-07-942-245-103	Sequence 103, App	197	4	36.4	104	2	US-08-273-146-63	Sequence 63, Appl
125	4	36.4	16	1	US-07-942-245-104	Sequence 104, App	198	4	36.4	104	2	US-08-727-688-32	Sequence 32, Appl
126	4	36.4	16	1	US-07-942-245-106	Sequence 106, App	199	4	36.4	105	1	US-08-422-101-9	Sequence 9, Appl1
127	4	36.4	16	1	US-07-942-245-110	Sequence 110, App	200	4	36.4	105	1	US-08-422-091-9	Sequence 9, Appl1
128	4	36.4	16	2	US-08-102-385C-35	Sequence 35, Appl	201	4	36.4	105	2	US-08-301-915-2	Sequence 2, Appl1
129	4	36.4	16	4	US-08-602-999A-365	Sequence 365, App	202	4	36.4	105	2	US-08-422-092-9	Sequence 9, Appl1
130	4	36.4	18	2	US-08-878-546-5	Sequence 365, App	203	4	36.4	105	2	US-08-788-800-6	Sequence 6, Appl1
131	4	36.4	19	2	US-08-878-546-1	Sequence 5, Appl1	204	4	36.4	105	3	US-08-524-694A-2	Sequence 2, Appl1
132	4	36.4	19	2	US-08-437-116A-2	Sequence 1, Appl1	205	4	36.4	105	3	US-08-422-093-9	Sequence 9, Appl1
133	4	36.4	19	2	US-08-437-116A-14	Sequence 14, Appl	206	4	36.4	105	3	US-08-422-112-9	Sequence 9, Appl1
134	4	36.4	19	2	US-08-530-792D-14	Sequence 14, Appl	207	4	36.4	106	1	US-08-488-113B-152	Sequence 152, App
135	4	36.4	19	2	US-08-530-792D-14	Sequence 14, Appl	208	4	36.4	106	1	US-08-477-484B-152	Sequence 152, App
136	4	36.4	20	1	US-07-956-848A-42	Sequence 42, Appl	209	4	36.4	106	2	US-08-546-560-152	Sequence 152, App
137	4	36.4	20	1	US-08-471-956-42	Sequence 42, Appl	210	4	36.4	106	3	US-08-776-404B-1	Sequence 1, Appl1
138	4	36.4	22	2	US-08-481-793-36	Sequence 36, Appl	211	4	36.4	106	4	US-08-839-765-152	Sequence 152, App
139	4	36.4	22	2	US-08-354-326-36	Sequence 36, Appl	212	4	36.4	106	4	US-09-136-389-152	Sequence 152, App
140	4	36.4	22	5	PCT-US95-07066B-36	Sequence 36, Appl	213	4	36.4	107	1	US-08-478-039-86	Sequence 86, Appl
141	4	36.4	23	4	US-09-282-351A-36	Sequence 36, Appl	214	4	36.4	107	1	US-08-399-106A-12	Sequence 12, Appl
142	4	36.4	26	1	US-07-942-245-331	Sequence 31, App	215	4	36.4	107	1	US-08-476-349A-86	Sequence 86, Appl
143	4	36.4	26	1	US-07-942-245-352	Sequence 352, App	216	4	36.4	107	1	US-08-433-105A-12	Sequence 12, Appl
144	4	36.4	26	1	US-07-942-245-378	Sequence 378, App	217	4	36.4	107	2	US-08-434-869A-12	Sequence 12, Appl
145	4	36.4	26	1	US-07-942-245-382	Sequence 382, App	218	4	36.4	109	1	US-07-745-382-22	Sequence 22, Appl
146	4	36.4	26	1	US-08-168-091A-24	Sequence 24, App	219	4	36.4	109	1	US-07-921-848-22	Sequence 22, Appl
147	4	36.4	29	1	US-08-367-198A-2	Sequence 2, Appl1	220	4	36.4	109	1	US-08-165-301A-22	Sequence 22, Appl
148	4	36.4	34	3	US-08-256-747C-49	Sequence 49, Appl	221	4	36.4	109	1	US-08-165-301A-26	Sequence 26, Appl
149	4	36.4	34	4	US-08-834-130A-49	Sequence 49, Appl	222	4	36.4	109	1	US-08-165-301A-28	Sequence 28, Appl
150	4	36.4	35	3	US-08-980-858A-34	Sequence 34, Appl	223	4	36.4	109	2	US-08-761-277A-51	Sequence 51, Appl
151	4	36.4	35	4	US-08-810-009-9	Sequence 9, Appl1	224	4	36.4	109	4	US-08-810-436-22	Sequence 22, Appl
152	4	36.4	35	4	US-08-810-009-10	Sequence 10, Appl	225	4	36.4	109	4	US-08-810-436-26	Sequence 26, Appl
153	4	36.4	35	4	US-08-810-009-11	Sequence 11, Appl	226	4	36.4	109	4	US-08-810-436-28	Sequence 28, Appl
154	4	36.4	35	4	US-08-810-009-12	Sequence 12, Appl	227	4	36.4	109	4	US-09-240-274-54	Sequence 54, Appl
155	4	36.4	35	4	US-08-810-009-13	Sequence 13, Appl	228	4	36.4	109	4	US-09-240-274-56	Sequence 56, Appl
156	4	36.4	37	1	US-08-673-269-7	Sequence 7, Appl1	229	4	36.4	109	4	US-09-240-274-57	Sequence 57, Appl
157	4	36.4	37	3	US-08-256-747C-71	Sequence 71, Appl	230	4	36.4	109	4	US-09-240-274-58	Sequence 58, Appl
158	4	36.4	37	3	US-09-192-048-19	Sequence 19, Appl	231	4	36.4	109	4	US-09-240-274-59	Sequence 59, Appl
159	4	36.4	37	4	US-08-834-130A-71	Sequence 71, Appl	232	4	36.4	109	4	US-09-240-274-61	Sequence 61, Appl
160	4	36.4	38	1	US-08-673-269-8	Sequence 8, Appl1	233	4	36.4	109	4	US-09-240-274-180	Sequence 180, App
161	4	36.4	43	3	US-08-310-912A-83	Sequence 83, Appl	234	4	36.4	109	5	PCT-US94-14179-22	Sequence 22, App
162	4	36.4	43	3	US-08-841-089-83	Sequence 83, Appl	235	4	36.4	109	5	PCT-US94-14179-26	Sequence 26, Appl
163	4	36.4	43	5	PCT-US95-04570-83	Sequence 83, Appl	236	4	36.4	109	5	PCT-US94-14179-28	Sequence 28, Appl
164	4	36.4	44	5	PCT-US95-04589-83	Sequence 83, Appl	237	4	36.4	110	1	US-08-199-311-2	Sequence 2, Appl1
165	4	36.4	44	1	US-08-405-392-11	Sequence 11, Appl	238	4	36.4	110	2	US-08-878-546-10	Sequence 10, Appl
166	4	36.4	44	3	US-08-487-691-11	Sequence 11, Appl	239	4	36.4	110	4	US-09-240-274-60	Sequence 60, Appl
167	4	36.4	45	2	US-08-687-916-19	Sequence 19, Appl	240	4	36.4	110	4	US-09-240-274-63	Sequence 63, Appl
168	4	36.4	45	4	US-09-138-614-19	Sequence 19, Appl	241	4	36.4	110	4	US-09-240-274-65	Sequence 65, Appl
169	4	36.4	54	4	US-08-905-223-427	Sequence 427, App	242	4	36.4	111	1	US-07-754-918A-11	Sequence 11, Appl
170	4	36.4	55	2	US-08-117-952-775	Sequence 775, App	243	4	36.4	111	1	US-08-264-093-6	Sequence 6, Appl1
171	4	36.4	58	4	US-08-927-319-39	Sequence 39, Appl	244	4	36.4	111	1	US-08-284-938B-16	Sequence 16, Appl
172	4	36.4	63	2	US-08-572-447C-9	Sequence 9, Appl1	245	4	36.4	111	1	US-08-284-938B-16	Sequence 16, Appl
173	4	36.4	68	2	US-08-117-952-771	Sequence 771, App	246	4	36.4	111	1	US-08-369-796-18	Sequence 18, Appl

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 15, 2001, 12:33:43 ; Search time 31.32 Seconds  
(without alignments)  
7.232 Million cell updates/sec

Title: US-09-372-036-29

Perfect score: 11

Sequence: 1 QOQTAPKAPTE 11

## Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

## Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTOS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	11	100.0	11	2	US-08-456-670B-29	Sequence 29, Appl
2	11	100.0	12	2	US-08-456-670B-42	Sequence 42, Appl
3	11	100.0	21	1	US-08-127-499A-33	Sequence 33, Appl
4	11	100.0	21	1	US-08-482-847-33	Sequence 33, Appl
5	11	100.0	232	2	US-08-456-670B-39	Sequence 39, Appl
6	11	100.0	478	2	US-08-456-670B-40	Sequence 40, Appl
7	11	100.0	484	1	US-08-127-499A-26	Sequence 26, Appl
8	11	100.0	484	1	US-08-482-847-26	Sequence 26, Appl
9	81.8	23	2	2	US-08-456-670B-20	Sequence 20, Appl
10	7	63.6	7	1	US-08-127-499A-34	Sequence 34, Appl
11	7	63.6	7	1	US-08-482-847-34	Sequence 34, Appl
12	6	54.5	890	1	US-08-145-006C-5	Sequence 5, Appl
13	6	54.5	890	1	PCT-US94-00545-5	Sequence 5, Appl
14	5	45.5	13	2	US-08-456-670B-36	Sequence 36, Appl
15	5	45.5	25	2	US-08-456-670B-23	Sequence 23, Appl
16	5	45.5	41	3	US-09-053-197A-46	Sequence 46, Appl
17	5	45.5	55	2	US-08-102-385G-15	Sequence 15, Appl
18	5	45.5	77	2	US-08-102-385G-29	Sequence 29, Appl
19	5	45.5	93	1	US-08-241-853-18	Sequence 18, Appl
20	5	45.5	93	1	US-08-850-917-18	Sequence 18, Appl
21	5	45.5	119	2	US-08-710-749-27	Sequence 27, Appl
22	5	45.5	135	4	US-08-905-223-464	Sequence 464, Appl
23	5	45.5	190	2	US-08-462-481-5	Sequence 5, Appl
24	5	45.5	190	2	US-08-436-771-7	Sequence 7, Appl
25	5	45.5	190	2	US-08-436-998-7	Sequence 7, Appl
26	5	45.5	190	2	US-08-487-797-7	Sequence 7, Appl
27	5	45.5	190	2	US-08-701-005A-5	Sequence 5, Appl

28	5	45.5	190	2	US-08-479-895-5	Sequence 5, Appl
29	5	45.5	190	5	PCT-US95-02058-7	Sequence 7, Appl
30	5	45.5	278	2	US-08-865-203-6	Sequence 6, Appl
31	5	45.5	278	2	US-07-849-420-6	Sequence 6, Appl
32	5	45.5	278	4	US-09-233-854-6	Sequence 6, Appl
33	5	45.5	279	2	US-09-135-658-4	Sequence 4, Appl
34	5	45.5	279	6	5472855-6	Patent No. 5472855
35	5	45.5	280	3	US-08-464-517-6	Sequence 6, Appl
36	5	45.5	280	3	US-08-463-772-6	Sequence 6, Appl
37	5	45.5	291	5	PCT-US93-05000-6	Sequence 6, Appl
38	5	45.5	292	2	US-08-464-517-23	Sequence 23, Appl
39	5	45.5	292	2	US-08-246-361A-6	Sequence 6, Appl
40	5	45.5	292	2	US-08-246-361A-23	Sequence 23, Appl
41	5	45.5	292	3	US-08-463-772-23	Sequence 23, Appl
42	5	45.5	292	5	PCT-US93-05000-23	Sequence 23, Appl
43	5	45.5	296	1	US-08-428-926-2	Sequence 2, Appl
44	5	45.5	296	1	US-08-435-434-5	Sequence 5, Appl
45	5	45.5	296	1	US-08-435-436-5	Sequence 5, Appl
46	5	45.5	296	1	US-08-428-927-2	Sequence 2, Appl
47	5	45.5	296	1	US-08-428-298-2	Sequence 2, Appl
48	5	45.5	296	1	US-08-339-517-2	Sequence 2, Appl
49	5	45.5	296	2	US-08-438-863-5	Sequence 5, Appl
50	5	45.5	296	3	US-08-986-217-6	Sequence 9, Appl
51	5	45.5	320	1	US-09-026-587-4	Sequence 4, Appl
52	5	45.5	320	1	US-08-726-525-4	Sequence 4, Appl
53	5	45.5	320	2	US-08-487-942-4	Sequence 4, Appl
54	5	45.5	320	2	US-08-726-036A-4	Sequence 4, Appl
55	5	45.5	343	2	US-08-446-875-2	Sequence 2, Appl
56	5	45.5	343	2	US-08-102-385G-2	Sequence 2, Appl
57	5	45.5	344	1	US-08-403-866-9	Sequence 9, Appl
58	5	45.5	378	2	US-08-986-217-6	Sequence 6, Appl
59	5	45.5	380	2	US-09-026-587-4	Sequence 4, Appl
60	5	45.5	380	2	US-09-227-420-4	Sequence 4, Appl
61	5	45.5	398	4	US-08-810-712-4	Sequence 4, Appl
62	5	45.5	438	2	US-08-458-346-14	Sequence 14, Appl
63	5	45.5	438	3	US-08-889-419-14	Sequence 14, Appl
64	5	45.5	438	1	PCT-US93-07189-3	Sequence 3, Appl
65	5	45.5	493	1	US-07-872-673B-3	Sequence 3, Appl
66	5	45.5	493	2	US-08-557-1122A-31	Sequence 31, Appl
67	5	45.5	508	2	US-08-557-1122A-37	Sequence 37, Appl
68	5	45.5	509	2	US-08-557-1122A-27	Sequence 27, Appl
69	5	45.5	510	2	US-08-557-1122A-29	Sequence 29, Appl
70	5	45.5	510	2	US-08-441-139-20	Sequence 20, Appl
71	5	45.5	510	2	US-08-557-1122A-28	Sequence 28, Appl
72	5	45.5	510	2	US-08-557-1122A-36	Sequence 30, Appl
73	5	45.5	520	2	US-09-091-432-2	Sequence 36, Appl
74	5	45.5	554	4	US-09-167-299-4	Sequence 4, Appl
75	5	45.5	746	5	PCT-US95-10509-2	Sequence 2, Appl
76	5	45.5	859	4	US-09-149-934-1	Sequence 1, Appl
77	5	45.5	1237	1	US-08-241-853-2	Sequence 2, Appl
78	5	45.5	1237	2	US-08-850-917-2	Sequence 2, Appl
79	5	45.5	1237	2	US-08-750-1122A-2	Sequence 2, Appl
80	5	45.5	1876	2	US-08-609-049A-12	Sequence 12, Appl
81	5	45.5	1876	2	US-08-609-049A-28	Sequence 28, Appl
82	5	45.5	2756	1	US-08-375-709-11	Sequence 11, Appl
83	5	45.5	2756	1	US-08-752-929-11	Sequence 11, Appl
84	5	45.5	2756	4	US-09-090-793-7	Sequence 7, Appl
85	5	45.5	3052	2	US-08-557-1122A-26	Sequence 26, Appl
86	5	45.5	364	4	US-08-218-027A-4	Sequence 4, Appl
87	5	45.5	7	1	US-08-218-027A-5	Sequence 5, Appl
88	5	45.5	7	1	US-08-211-747-5	Sequence 5, Appl
89	5	45.5	7	2	US-08-340-283-183	Sequence 183, App
90	5	45.5	9	2	US-08-340-283-195	Sequence 195, App
91	5	45.5	9	2	US-08-456-670B-25	Sequence 25, Appl
92	5	45.5	10	3	PCT-US94-03796-5	Sequence 5, Appl
93	5	45.5	364	4	US-08-822-586-8	Sequence 8, Appl
94	5	45.5	364	4	US-08-822-586-12	Sequence 12, Appl
95	5	45.5	11	2	US-08-456-670B-38	Sequence 38, Appl
96	5	45.5	12	1	US-08-196-940-1	Sequence 1, Appl
97	5	45.5	12	1	US-08-211-747-2	Sequence 1, Appl
98	5	45.5	12	2	US-08-666-473-113	Sequence 113, App
99	5	45.5	12	4	US-08-602-999A-252	Sequence 252, App
100	5	45.5	12	4	US-08-960-780-61	Sequence 61, Appl

**THIS PAGE BLANK (uspto)**

---

OS Pneumocystis carinii f. sp. ratti.  
 OC Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes; Pneumocystidaceae;  
 OC Pneumocystis.  
 OX NCBI\_TaxID=38082;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20136152; PubMed=10669586;  
 RA Schaffzin J.K., Garbe T.R., Stringer J.R.;  
 RT "Major surface glycoprotein genes from pneumocystis carinii f. sp. ratti."  
 RL Fungal Genet. Biol. 28:214-226(1999).  
 DR EMBL; AF184210; AAF63434.1; -.  
 DR InterPro; IPR003330; -.  
 DR Pfam; PF02349; MSG; 1.  
 FT NON\_TER 1 1  
 FT SMART; IPR000970; -.  
 SO SEQUENCE 120 AA; 12879 MW; 04EF193EFF2079A9 CRC64;

Query Match 45.5%; Score 5; DB 3; Length 120;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 DB 9 APKAP 13

RESULT 49  
 O9PP33 PRELIMINARY; PRT; 121 AA.  
 AC O9PP33;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE HIT-FAMILY PROTEIN.  
 GN C30898.  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 OC Campylobacter.  
 OX NCBI\_TaxID=197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCITC 11168;  
 RX MEDLINE=20150912; PubMed=10688204;  
 RA Parkhill J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,  
 RA Basham D., Chillingworth T., Davies K.M., Feltham T., Holtroyd S.,  
 RA Jagsels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
 RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,  
 RA Whitehead S., Barrall B.G.;  
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences."  
 RL Nature 403:665-668(2000).  
 DR EMBL; AL139076; CAB73156.1; -.  
 DR InterPro; IPR001310; -.  
 DR Pfam; PF01230; HIT; 1.  
 DR PRINTS; PR00332; HISTRIAD.  
 DR PROSITE; PS00892; HIT; UNKNOWN\_1.  
 SO SEQUENCE 121 AA; 13681 MW; 81B19E6CB53E9DD CRC64;

Query Match 45.5%; Score 5; DB 2; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 86;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 DB 33 APKAP 37

RESULT 50  
 O9PS90 PRELIMINARY; PRT; 124 AA.  
 ID O9PS90

AC O9PS90;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE WNT-3A PROTEIN (FRAGMENT).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96076124; PubMed=757679;  
 RA Holliday M., McMahon J.A., McMahon A.P.;  
 RT "Wnt expression patterns in chick embryo nervous system."  
 RL Mech. Dev. 52:9-25(1995).  
 DR InterPro; IPR000970; -.  
 DR Pfam; PF00110; WNT; 1.  
 DR SMART; SM00097; WNT1; 1.  
 SO SEQUENCE 124 AA; 14291 MW; 21E6C46AE7E2760B CRC64;

Query Match 45.5%; Score 5; DB 13; Length 124;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 KAPTE 11  
 |||||  
 DB 58 KAPTE 62

Search completed: August 15, 2001, 12:42:45  
 Job time: 530 sec



GN CG14554.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,  
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,  
 RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Crawley S., Dahle C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A.A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Sliker E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003754; AAF56570.1; -;  
 DR FlyBase: FBgn0039412; CG14554.  
 SO SEQUENCE 114 AA; 13638 MW; C649FD265FDC5CFE CRC64;

Query Match 45.5%; Score 5; DB 5; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 82;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOOTA 5  
 |||||  
 DB 105 QOOTA 109

RESULT 46  
 ID 09NLN7 PRELIMINARY; PRT; 117 AA.  
 AC 09NLN7;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DE POSSIBLE L7138-8-RELATED PROTEIN (FRAGMENT).  
 GN LM26.405.  
 OS Leishmania major.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI\_TaxID=5664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-FRIEDLIN;  
 RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL160493; CAB98023.1; -;  
 FT NON\_TER 117  
 FT 117  
 SO SEQUENCE 117 AA; 12967 MW; 9DE8EBDEE514C372 CRC64;

Query Match 45.5%; Score 5; DB 5; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 84;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOOTA 5  
 |||||  
 DB 48 QOOTA 52

RESULT 47  
 ID 09RTX9 PRELIMINARY; PRT; 118 AA.  
 AC 09RTX9;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DE HT FAMILY PROTEIN.  
 GN DR1621.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
 NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RI;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Oln H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 radiodurans RI.";  
 RL Science 286:1571-1577(1999).  
 DR EMBL: AE002005; AAF11181.1; -;  
 DR HSSP: P49773; IKPF.  
 DR TIGR: DR1621; -;  
 DR InterPro: IPR001310; -;  
 DR Pfam: PF01230; HIT: 1;  
 DR PRINTS: PRO0332; HISTRIAD.  
 DR PROSITE: PS00892; HIT: 1.  
 SO SEQUENCE 118 AA; 13112 MW; AA6C3A4FC73EF912 CRC64;

Query Match 45.5%; Score 5; DB 2; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
 |||||  
 DB 40 APRAP 44

RESULT 48  
 ID 09PBW7 PRELIMINARY; PRT; 120 AA.  
 AC 09PBW7;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE MAJOR SURFACE GLYCOPROTEIN (FRAGMENT).

RC STRAIN-A3(2); PubMed=8843436;  
 RA Redenbach M., Kieser H.M., Denaplatte D., Eichner A., Cullum J.,  
 RA Khashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL163641; CAB87228.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 103 AA; 10639 MW; 9316F1F38C1BD492 CRC64;

Query Match 45.5%; Score 5; DB 2; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OQOTA 5  
 |||||  
 DB 40 OQOTA 44

## RESULT 42

O9N3V6 PRELIMINARY; PRT; 110 AA.  
 AC O9N3V6;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE HYPOTHEICAL PROTEIN Y46E12A.D.  
 GN Y46E12A.D.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodermidae; Caenorhabditis.  
 NCBI\_TaxID=6239;  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC024788; AAF60614.1; -  
 SQ SEQUENCE 110 AA; 12150 MW; BC5FBF808B497E66 CRC64;

Query Match 45.5%; Score 5; DB 5; Length 110;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 DB 20 APKAP 24

## RESULT 43

O9ESAB PRELIMINARY; PRT; 111 AA.  
 AC O9ESAB;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE SMPF NEUREGULIN BETA 2 (FRAGMENT).  
 GN NRCL.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY;  
 RC TISSUE=AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;  
 RA Carroll S.L., Anderson K.D., Frohner P.W.;  
 RT "Structural and Functional Diversity of SMD Neuregulin Splice  
 RT Variants Expressed in the Adult Rat Nervous System.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF194441; AAG28430.1; -  
 FT NON\_TER 1  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 111 AA; 12198 MW; CC8BB870584C9F8C CRC64;

Query Match 45.5%; Score 5; DB 11; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QAPK 7  
 |||||  
 DB 38 QAPK 42

## RESULT 44

O95014 PRELIMINARY; PRT; 114 AA.  
 AC O95014;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 DE WUGSC:H.DJ0855D21.2 PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 NCBI\_TaxID=9606;  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Courtney L., Langston Y., Stoneking T., Drone K., Shih K.;  
 RT "The sequence of Homo sapiens PAC clone DJ0855D21.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Waterston R.H.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Waterston R.H.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Waterston R.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC004908; AAD05196.1; -  
 SQ SEQUENCE 114 AA; 12265 MW; 1455C818AE3FB962 CRC64;

Query Match 45.5%; Score 5; DB 4; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 82;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
 |||||  
 DB 36 TAPKA 40

## RESULT 45

O9VBH2 PRELIMINARY; PRT; 114 AA.  
 AC O9VBH2;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE CG14554 PROTEIN.

DR PRINTS: PR00835; ROPREGULATRY.  
 KW Plasmid.  
 SO SEQUENCE 64 AA; 7298 MW; 514D070E518CF78F CRC64;

Query Match 45.5%; Score 5; DB 2; Length 64;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 OQORA 5  
 |||||  
 Db 4 OQORA 8

RESULT 38  
 O9MAC9 PRELIMINARY; PRT; 82 AA.  
 AC O9MAC9;  
 DT 01-OCT-2000 (TREMUREL. 15, Created)  
 DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMUREL. 16, Last annotation update)  
 DE HYPOHETICAL 8.9 KDA. PROTEIN (FRAGMENT).  
 GN BCI-7.  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae;  
 OC Hordeum.  
 OC NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. INGRID; TISSUE=LEAF;  
 RA Besser K., Jarosch B., Langen G., Kogel K.H.;  
 RT \*Identification and expression analysis of genes induced in barley by  
 RT chemical that activate disease resistance."  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ250663; CAB71340.1; -  
 DR InterPro: IPR000864; -  
 DR Pfam: PF00280; potato\_inhib1; 1.  
 DR PRINTS: PR00292; POTATOINHTR.  
 DR Hypothetical protein.  
 KW NON\_TER 1  
 FT SEQUENCE 82 AA; 8862 MW; B13F86FBEB78AE9 CRC64;  
 SO

Query Match 45.5%; Score 5; DB 10; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 Db 13 APKAP 17

RESULT 39  
 O9YGP4 PRELIMINARY; PRT; 88 AA.  
 AC O9YGP4;  
 DT 01-MAY-1999 (TREMUREL. 10, Created)  
 DT 01-NOV-1999 (TREMUREL. 12, Last sequence update)  
 DT 01-MAY-2000 (TREMUREL. 13, Last annotation update)  
 DE WMT-3A HOMOLOG (FRAGMENT).  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Huges S.M., Crnogorac-Jurcevic T., Sillibourne J., Bunnstead N.;  
 RT Assignment of a close homologue of murine wingless-type WMTV  
 RT Integration site family member 3a (WNT3A) to chicken chromosome 2."  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF109144; AAD19652.2; -

DR InterPro: IPR000970; -  
 DR Pfam: PF00110; wnt; 1.  
 FT NON\_TER 1  
 FT NON\_TER 88  
 SO SEQUENCE 88 AA; 10322 MW; 9124F1CGAFB7F108 CRC64;

Query Match 45.5%; Score 5; DB 13; Length 88;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 7 KAPTE 11  
 |||||  
 Db 56 KAPTE 60

RESULT 40  
 O38792 PRELIMINARY; PRT; 95 AA.  
 AC O38792;  
 DT 01-NOV-1996 (TREMUREL. 01, Created)  
 DT 01-NOV-1996 (TREMUREL. 01, Last sequence update)  
 DT 01-AUG-1998 (TREMUREL. 07, Last annotation update)  
 DE HYBRID OF A. MAJUS AND A. HISPANICUM PTL2 MRNA.  
 GN PTL2.  
 OS Antirrhinum sp. (snapdragon).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
 OC Lamiales; Scrophulariaceae; Antirrhinum.  
 OC NCBI\_TaxID=29720;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SRRAIN-7249 17; TISSUE=STYLAR TISSUE;  
 RA Baldwin T.C., Coen E.S., Dickinson H.G.;  
 RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: X71782; CAA50667.1; -  
 SO SEQUENCE 95 AA; 9964 MW; 96C39D978C7821BF CRC64;

Query Match 45.5%; Score 5; DB 10; Length 95;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 Db 63 APKAP 67

RESULT 41  
 O9L063 PRELIMINARY; PRT; 103 AA.  
 AC O9L063;  
 DT 01-OCT-2000 (TREMUREL. 15, Created)  
 DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMUREL. 15, Last annotation update)  
 DE HYPOHETICAL 10.6 KDA. PROTEIN.  
 GN SCC105.22C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OC NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SRRAIN-A3(2);  
 RA Brown S.P., Harris D.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SRRAIN-A3(2);  
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.

RT amplification.";  
 RL Can. J. Microbiol. 39:402-411(1993).  
 CC -1- FUNCTION: PARTICIPATES IN SULFATE RESPIRATION COUPLED WITH  
 CC PHOSPHORYLATION BY TRANSFERRING ELECTRONS FROM THE ENZYME  
 CC DEHYDROGENASE TO FERREDOXIN.  
 CC -1- PFM: BINDS FOUR NONPARALLEL HEME GROUPS PER MOLECULE.  
 CC -1- SIMILARITY: TO OTHER C3-TYPE CYTOCHROMES.  
 DR HSP: P00132; 2CDV.  
 KM Electron transport; Sulfate respiration; Heme.  
 FT NON\_TER 1 1  
 FT METAL 13 13 IRON (HEME 1 AXIAL LIGAND) (BY  
 FT METAL 16 16 SIMILARITY).  
 FT METAL 16 16 IRON (HEME 3 AXIAL LIGAND) (BY  
 FT METAL 16 16 SIMILARITY).  
 FT NON\_TER 23 23  
 FT SEQUENCE 23 AA; 2477 MW; 1114D6AC22FAD6B CRC64;

Query Match 45.5%; Score 5; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 AKAP 9  
 DB 2 AKAP 6

RESULT 35  
 O56859 PRELIMINARY; PRT; 61 AA.  
 AC O56859;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE ADENYLATE KINASE (FRAGMENT).  
 GN ADK.  
 OS Yersinia enterocolitica.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OX NCBI\_Taxid=630;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=8081-C;  
 RX MEDLINE=97086507; Pubmed=8932701;  
 RA Zhang L., Tolvanen P., Skurnik M.;  
 RT "The gene cluster directing O-antigen biosynthesis in Yersinia  
 RT enterocolitica serotype O:8: Identification of the genes for mannose  
 RT and galactose biosynthesis and the gene for the O-antigen  
 RT polymerase.";  
 RL Microbiology 142:277-288(1996).  
 DR EMBL: U46859; AAC60759.1; -.  
 DR HSP: P05082; LAKE.  
 FT NON\_TER 1 1  
 FT SEQUENCE 61 AA; 6941 MW; 68407A9B73F0DD69 CRC64;

Query Match 45.5%; Score 5; DB 2; Length 61;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OOTAP 6  
 DB 10 OOTAP 14

RESULT 36  
 O9S1E1 PRELIMINARY; PRT; 64 AA.  
 AC O9S1E1;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE PUTATIVE REPLICATION REGULATORY PROTEIN.

GN YPPCp1.03.  
 OS Yersinia pestis.  
 OG Plasmid pPCp1.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Yersinia.  
 OX NCBI\_Taxid=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92 BIOVAR ORIENTALIS;  
 RA Baker S.G., Mungall K.;  
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92 BIOVAR ORIENTALIS;  
 RA Parhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92 BIOVAR ORIENTALIS;  
 RA Kariyshev A.V., Wren B.W.;  
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AL109969; CAB53166.1; -.  
 DR HSP: P03051; IROP.  
 DR InterPro: IPR000769; -.  
 DR Pfam: PF01815; Rop; 1.  
 DR PRINTS: PR00835; ROPREGULATRY.  
 DR ProDom: PD012167; -; 1.  
 KW Plasmid.  
 SQ SEQUENCE 64 AA; 7280 MW; 42D0A3562BA452FF CRC64;

Query Match 45.5%; Score 5; DB 2; Length 64;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OOOTTA 5  
 DB 4 OOOTTA 8

RESULT 37  
 O9JP69 PRELIMINARY; PRT; 64 AA.  
 AC O9JP69;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE ROM-LIKE PROTEIN.  
 GN ROM.  
 OS Escherichia coli.  
 OG Plasmid pEC156.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_Taxid=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EI585-68;  
 RX MEDLINE=99077830; Pubmed=98586694;  
 RA Kaczorowski T., Szybalski W.;  
 RT "Genomic DNA sequencing by SP6L-6 primer walking using hexamer  
 RT ligation.";  
 RL Gene 223:83-91(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EI585-68;  
 RA Biedrzycka I., Kaczorowska A., Sektas M., Kaczorowski T.;  
 RT "Characterization of pEC156, a ColEI-like plasmid from Escherichia  
 RT coli EI585-68 that carries the genes of the EcoVIII restriction-  
 RT modification system.";  
 RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF158026; AAF71175.1; -.  
 DR InterPro: IPR000769; -.  
 DR Pfam: PF01815; Rop; 1.

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarly L., Morris J., Mostreli A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003497; AAF48411.1; -;  
 DR FlyBase: FBgn0030600; h1w; -;  
 DR InterPro: IPR000315; -;  
 DR InterPro: IPR001298; -;  
 DR InterPro: IPR001841; -;  
 DR SMART: SM00336; BBOX: 1.  
 SO SEQUENCE 2478 AA; 270371 MW; 727D854E6C51B3FA CRC64;

Query Match 54.5%; Score 6; DB 5; Length 2478;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOQTAP 6  
 |||||  
 Db 1736 QOQTAP 1741

RESULT 32  
 ID 09NB71 PRELIMINARY; PRT; 5233 AA.  
 AC 09NB71;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE HIGHWIRE  
 GN h1w OR CG9041 OR CG9049.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; *Drosophila*.  
 NC NCBL\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CANTON S;  
 RX MEDLINE=20296162; PubMed=10839352;  
 RA Wen H.I., Dianzono A., Fetter R.D., Bergstrom K., Strauss R.,  
 RA Goodman C.S.;  
 RT "Highwire regulates synaptic growth in *Drosophila*.";  
 RL Neuron 26:313-329(2000).  
 DR EMBL: AF262977; AAF76150.1; -;  
 DR FlyBase: FBgn0030600; h1w; -;  
 DR InterPro: IPR000315; -;  
 DR InterPro: IPR000345; -;  
 DR InterPro: IPR000408; -;  
 DR InterPro: IPR001298; -;  
 DR InterPro: IPR001841; -;  
 DR Pfam: PF00415; RCC1; 4.  
 DR PRINTS: PR00633; RCCNSATON.  
 DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 DR PROSITE: PS00626; RCL1\_2; UNKNOWN\_2.  
 DR SMART: SM00336; BBOX: 1.

SQ SEQUENCE 5233 AA; 565692 MW; 47141BF02CE07E25 CRC64;

Query Match 54.5%; Score 6; DB 5; Length 5233;  
 Best Local Similarity 100.0%; Pred. No. 1,4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOQTAP 6  
 |||||  
 Db 4491 QOQTAP 4496

RESULT 33  
 ID P81150 PRELIMINARY; PRT; 22 AA.  
 AC P81150;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)  
 DE CYTOCHROME C3 (FRAGMENT).  
 OS *Desulfovibrio vulgaris*.  
 OC Bacteria; Proteobacteria; delta subdivision; *Desulfovibrio*.  
 NC NCBL\_TaxID=881;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93272123; PubMed=8388770;  
 RA Kwon D.Y., Vedvick T.S., McCue A.F., Gevertz D.;  
 RT "Rapid comparison of the cytochrome c3 gene from nine strains of  
 RT *Desulfovibrio vulgaris* using polymerase chain reaction  
 RT amplification.";  
 RL Can. J. Microbiol. 39:402-411(1993).  
 CC -1- FUNCTION: PARTICIPATES IN SULFATE RESPIRATION COUPLED WITH  
 CC PHOSPHORYLATION BY TRANSFERRING ELECTRONS FROM THE ENZYME  
 CC DEHYDROGENASE TO FERREDOXIN.  
 CC -1- PTM: BINDS FOUR NONPARALLEL HEME GROUPS PER MOLECULE.  
 CC -1- SIMILARITY: TO OTHER C3-TYPE CYTOCHROMES.  
 DR HSSP: P00131; 2CTH.  
 KM Electron transport; Sulfate respiration; Heme.  
 FT NON\_TER 1 1  
 FT METAL 12 12 IRON (HEME 1 AXIAL LIGAND) (BY  
 FT METAL 15 15 SIMILARITY).  
 FT IRON (HEME 3 AXIAL LIGAND) (BY  
 FT SIMILARITY).  
 FT NON\_TER 22 22  
 SO SEQUENCE 22 AA; 2350 MW; 501C75E1C2225A6C CRC64;

Query Match 45.5%; Score 5; DB 2; Length 22;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
 |||||  
 Db 1 APRAP 5

RESULT 34  
 ID P81149 PRELIMINARY; PRT; 23 AA.  
 AC P81149;  
 DT 01-JAN-1998 (TReMBLrel. 05, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)  
 DE CYTOCHROME C3 (FRAGMENT).  
 OS *Desulfovibrio vulgaris*.  
 OC Bacteria; Proteobacteria; delta subdivision; *Desulfovibrio*.  
 NC NCBL\_TaxID=881;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=93272123; PubMed=8388770;  
 RA Kwon D.Y., Vedvick T.S., McCue A.F., Gevertz D.;  
 RT "Rapid comparison of the cytochrome c3 gene from nine strains of  
 RT *Desulfovibrio vulgaris* using polymerase chain reaction

```

RP SEQUENCE FROM N.A.
RA Anderson J.B., Yamashita R.A., Sellers J.R.;
RT "Complete cDNA for an unconventional Myosin (Class VII) in Drosophila
RT melanogaster."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF233269; AAF34810.1; -.
DR FlyBase; FBgn0040299; Myo28B1.
DR InterPro; IPR000048; -.
DR InterPro; IPR000299; -.
DR InterPro; IPR000857; -.
DR InterPro; IPR001609; -.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00612; IQ; 3.
DR Pfam; PF00784; MYTH4; 2.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PROSITE; PS50057; BAND_41_3; 2.
DR SMART; SM00295; B41; 1.
SO SEQUENCE 2121 AA; 243312 MW; 1D2ED04C8D1DA7 CRC64;

Query Match          54.5%; Score 6; DB 5; Length 2121;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAPT 10
DB 915 APKAPT 920

RESULT 30
O9VL23 PRELIMINARY; PRT; 2129 AA.
AC O9VL23;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CG6976 PROTEIN.
GN MYO28B1 OR CG6976.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
PC STRAIN-BERKELEY.
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Butkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA Deodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
RA Jaiswal M., Katush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Meltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mervulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

```

```

RA Palazolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003618; AAF52536.1; -.
DR HSSP; P08799; 1MND.
DR FlyBase; FBgn0040299; Myo28B1.
DR InterPro; IPR000048; -.
DR InterPro; IPR000299; -.
DR InterPro; IPR000857; -.
DR InterPro; IPR001609; -.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00612; IQ; 3.
DR Pfam; PF00784; MYTH4; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; -.
DR PROSITE; PS50057; BAND_41_3; 2.
DR SMART; SM00295; B41; 1.
SO SEQUENCE 2129 AA; 244144 MW; 1A90EBDC35103481 CRC64;

Query Match          54.5%; Score 6; DB 5; Length 2129;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAPT 10
DB 921 APKAPT 926

RESULT 31
O9VX25 PRELIMINARY; PRT; 2478 AA.
AC O9VX25;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE CG9041 PROTEIN.
GN HIW OR CG9041 OR CG9049.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
[1]
PC STRAIN-BERKELEY.
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Butkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA Deodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abrell J.F., Agbayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,  
 RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ikegawa C.,  
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spiller E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Welschenbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AEO03831; AAF58849.1;  
 DR FlyBase: FBgn003482; CG1371.  
 SQ SEQUENCE 1199 AA; 130890 MW; 043D72C8D2D5E71F CRC64;

Query Match 54.5%; Score 6; DB 5; Length 1199;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 6 PKAPTE 11  
 Db 597 PKAPTE 602  
 RESULT 27  
 013617 PRELIMINARY; PRT; 1241 AA.  
 AC 013617;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE TAT-BINDING HOMOLOG 7.  
 GN P1026.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes;  
 OX NCBI\_TaxID=4896;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-972 H-;  
 RA Kushida N., Yamazaki S., Tanaka T., Jinno K., Halkawa Y., Yamazaki J.,  
 RA Yamamoto S., Sekine M., Oguchi A., Nagai Y., Sakai K.,  
 RA Ogura K., Otsuka R., Kudo Y., Yanagida M., Machida M., Zhang M.Q.,  
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB004535; BAA21405.1;  
 DR InterPro: IPR001939;  
 DR InterPro: IPR003593;  
 DR Pfam: PF00004; AAA; 1.

DR PROSITE: PS00674; AAA; 1.  
 DR SMART: SM00382; AAA; 1.  
 SQ SEQUENCE 1241 AA; 142188 MW; F19BD7AF49BF8755 CRC64;  
 Query Match 54.5%; Score 6; DB 3; Length 1241;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 6 PKAPTE 11  
 Db 202 PKAPTE 207  
 RESULT 28  
 09NF04 PRELIMINARY; PRT; 1257 AA.  
 AC 09NF04;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE Y1058C.B. PROTEIN.  
 GN Y1058C.B.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderiinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sulston J.E.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE:99069613; PubMed-9851916;  
 RA none;  
 RT "Genome sequence of the nematode *C. elegans*: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL: AL132879; CAB60860.2;  
 DR InterPro: IPR000619;  
 DR InterPro: IPR000906;  
 DR InterPro: IPR001478;  
 DR Pfam: PF00595; PD2; 3.  
 DR Pfam: PF00791; ZU5; 1.  
 DR PROSITE: PSS0052; GUANYLATE\_KINASE\_2; 1.  
 DR SMART: SM00872; GuKc; 1.  
 SQ SEQUENCE 1257 AA; 139391 MW; 617E729417136819 CRC64;

Query Match 54.5%; Score 6; DB 5; Length 1257;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 QOQTAP 6  
 Db 1022 QOQTAP 1027  
 RESULT 29  
 09NH54 PRELIMINARY; PRT; 2121 AA.  
 AC 09NH54;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE MYO28B1 OR CG6976.  
 GN MYO28B1 OR CG6976.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Ephydrophillidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 [1]

OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 CC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Percy C.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latrelle P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 CC -1- SIMILARITY: TO DEAD/DEAH BOX HELICASE FAMILY.  
 CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.  
 CC -1- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES.  
 DR EMBL: 281094; CAB03153.1; -;  
 DR InterPro: IPR000629; -;  
 DR InterPro: IPR001410; -;  
 DR InterPro: IPR001650; -;  
 DR Pfam: PF00270; DEAD; 1.  
 DR Pfam: PF00271; helicase.C; 1.  
 DR PROSITE: PS00039; DEAD\_ATP\_HELICASE; 1.  
 DR SMART: SM00490; HELICC; 1.  
 DR AT-Blinding; Helicase; RNA-binding.  
 KW SEQUENCE 995 AA; 107563 MW; 94A25950B93C93F5 CRC64;  
 SO

Query Match 54.5%; Score 6; DB 5; Length 995;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAPT 10  
 |||||  
 DB 889 APKAPT 894

RESULT 24  
 ID 088532 PRELIMINARY; PRT: 1052 AA.  
 AC 088532;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE ZINC FINGER RNA BINDING PROTEIN.  
 GN ZFR.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=TESTIS;  
 RA Mesgher M.J., Schumacher J.M., Lee K., Holdcraft R.W., Edelhoff S.,  
 RA Distche C., Braun R.E.;  
 RT "Identification of ZFR, an ancient and highly conserved chromosome-  
 associated zinc finger RNA binding protein.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF011059; AAC25762.1; -;  
 DR MGD: MGI:1341890; Zfr.  
 DR InterPro: IPR000690; -;  
 DR InterPro: IPR000822; -;

DR Pfam: PF00096; Zf-C2H2; 3.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; UNKNOWN\_3.  
 DR SMART: SM00355; Znf\_C2H2; 1.  
 KW DNA-binding; Zinc-finger.  
 SO SEQUENCE 1052 AA; 114434 MW; 29E2927CC0B60B9D CRC64;  
 SO

Query Match 54.5%; Score 6; DB 11; Length 1052;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QTPAKA 8  
 |||||  
 DB 165 QTPAKA 170

RESULT 25  
 ID 097169 PRELIMINARY; PRT: 1158 AA.  
 AC 097169;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE HAIRLESS PROTEIN.  
 GN HAIRLESS.  
 OS Drosophila hydei (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7224;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=WILD TYPE; CV: HESS; O.;  
 RA Marquart J., Alexief-Damianof C., Preiss A., Maier D.;  
 RT "Rapid divergence in the course of Drosophila evolution reveals  
 structural important domains of the Notch antagonist Hairless.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ232783; CAB38221.1; -;  
 DR FLYBase: FBgn0019190; DhydVH.  
 SO SEQUENCE 1158 AA; 119635 MW; 0B611D3739CF397F CRC64;  
 SO

Query Match 54.5%; Score 6; DB 5; Length 1158;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKAP 9  
 |||||  
 DB 159 TAPKAP 164

RESULT 26  
 ID 09V5G1 PRELIMINARY; PRT: 1199 AA.  
 AC 09V5G1;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE CG1371 PROTEIN.  
 GN CG1371.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,



RL Science 282:2012-2018(1998).  
 DR EMBL: AL132889; CAB81909.1; -.  
 DR InterPro: IPR001680; -.  
 DR InterPro: IPR001865; -.  
 DR InterPro: IPR001899; -.  
 DR Pfam: PF00400; WD40; 8.  
 DR PROSITE: PS00343; GRAM\_POS\_ANCHORING; UNKNOWN\_1.  
 DR PROSITE: PS00962; RIBOSOMAL\_S2\_1; UNKNOWN\_1.  
 SO SEQUENCE 855 AA; 94259 MW; E4426D383235E008 CRC64;

Query Match 54.5%; Score 6; DB 5; Length 855;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 OTAPKA 8  
 Db 767 OTAPKA 772

## RESULT 21

O9HUA6 PRELIMINARY; PRT; 861 AA.  
 AC O9HUA6;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE PERIPLASMIC GLUCANS BIOSYNTHESIS PROTEIN MDH.  
 GN MDH OR PA5077.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 ON NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PA01  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 DR EMBL: AE004820; AAC08462.1; -.  
 DR InterPro: IPR001173; -.  
 DR Pfam: PF00535; Glycosyltransf\_2; 1.  
 DR SMART: SM00355; Glycosyltransf\_2; 1.  
 SO SEQUENCE 861 AA; 96994 MW; 63557B5A83107E7E CRC64;

Query Match 54.5%; Score 6; DB 2; Length 861;  
 Best Local Similarity 100.0%; Pred. No. 32;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 OTAPKA 8  
 Db 376 OTAPKA 381

## RESULT 22

O9VEM6 PRELIMINARY; PRT; 925 AA.  
 AC O9VEM6;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE CG10309 PROTEIN.  
 GN CG10309.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.  
 ON NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Cealiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,  
 RA April J.F., Abdyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Balling R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.A., Bouck J., Brokstein P., Brotler P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshirel A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacible J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003713; AAF5301.1; -.  
 DR HSSP: P08046; 1ALJ.  
 DR FlyBase: FBgn0038418; CG10309.  
 DR InterPro: IPR000822; -.  
 DR Pfam: PF00096; zf-C2H2; 4.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2; 4.  
 DR SMART: SM00355; ZNF\_C2H2; 1.  
 DR DNA-binding; Metal-binding; Zinc-finger.  
 SO SEQUENCE 925 AA; 100378 MW; E3866D4BCE2A2AF2 CRC64;

Query Match 54.5%; Score 6; DB 5; Length 925;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 OTAPKA 8  
 Db 618 OTAPKA 623

## RESULT 23

P90897 PRELIMINARY; PRT; 995 AA.  
 ID P90897;  
 AC P90897;  
 DT 01-MAY-1997 (TREMBlrel. 03, Created)  
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE F58G11.2 PROTEIN.  
 GN F58G11.2.

RESULT 17  
 ID Q9P196 PRELIMINARY; PRT: 663 AA.  
 AC Q9P196;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE FUMARATE REDUCTASE FLAVOPROTEIN SUBUNIT (EC 1.3.99.1).  
 GN FRDA OR CJD0409.  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 OC Campylobacter.  
 RX NCBI\_TaxID=197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCYC 11168;  
 RX MEDLINE=20150912; PubMed=10688204;  
 RA Parkhill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,  
 Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
 Jagsels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.M.,  
 Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,  
 Whitehead S., Barrell B.G.;  
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
 reveals hypervariable sequences.";  
 RL Nature 403:665-668(2000).  
 DR EMBL: AL139075; CAB74245.1; -  
 DR InterPro: IPR000464; -  
 DR Pfam: PF00890; FAD\_binding\_2; 1.  
 SQ SEQUENCE 663 AA: 73788 MW; 4FE340FEE783E9F9 CRC64;

Query Match 54.5%; Score 6; DB 2; Length 663;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 QTAAPK 8  
 DB 91 QTAAPK 96

RESULT 18  
 ID Q25586 PRELIMINARY; PRT: 672 AA.  
 AC Q25586;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE PARAMYOSIN-RELATED PROTEIN (FRAGMENT).  
 OS Onchocerca gibsoni.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
 OC Onchocercidae; Onchocerca.  
 RX NCBI\_TaxID=6284;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96360163; PubMed=8719949;  
 RA Zhang D., Miller D.J.;  
 RT "Characterization of a novel non-muscle myosin-related protein from  
 Onchocerca gibsoni.";  
 RL Int. J. Parasitol. 25:1385-1391(1995).  
 DR EMBL: U20609; AB50272.1; -  
 DR InterPro: IPR000727; -  
 FT NON\_TER 1  
 FT NON\_TER 672  
 SQ SEQUENCE 672 AA: 80074 MW; C610249CEFF0FFFD1 CRC64;

Query Match 54.5%; Score 6; DB 5; Length 672;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QOQTAP 6  
 DB 111111

DB 240 QOQTAP 245  
 RESULT 19  
 ID Q9RKV7 PRELIMINARY; PRT: 689 AA.  
 AC Q9RKV7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE PUTATIVE SERINE/THREONINE PROTEIN KINASE (FRAGMENT).  
 GN SC961.09.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomyces.  
 RX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Oliver K., Harris D.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kleser H.M., Denapalte D., Elchner A., Cullum J.,  
 Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL132997; CAB61321.1; -  
 KW Kinase; Serine/threonine-protein kinase.  
 FT NON\_TER 689  
 FT NON\_TER 689  
 SQ SEQUENCE 689 AA: 73645 MW; 6CF7BF01BC8556DA CRC64;

Query Match 54.5%; Score 6; DB 2; Length 689;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 TAPKAP 9  
 DB 659 TAPKAP 664

RESULT 20  
 ID Q9NEW7 PRELIMINARY; PRT: 855 AA.  
 AC Q9NEW7;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE Y11B2G.H PROTEIN.  
 GN Y11B2G.H.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 RX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sulston J.E.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9905613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 investigating biology.";

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshneft A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spler E., Spredling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Glibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003608; AAF52199.1; -;  
 DR FLYBase: FBgn0031667; CG18623.  
 DR InterPro: IPR002048; -;  
 DR InterPro: IPR002965; -;  
 DR PRINTS: PRO1217; PRICHEXTENSN.  
 DR PROSITE: PS00018; EF\_HAND; UNKNOWN.1.  
 SO SEQUENCE 388 AA; 40889 MW; 13903464EBC79CB3 CRC64;

Query Match 54.58; Score 6; DB 5; Length 388;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAPT 10  
 |||||  
 DB 325 APRAPT 330

RESULT 15  
 ID 086584 PRELIMINARY; PRT; 606 AA.  
 AC 086584;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE PUTATIVE ATP/GTP BINDING PROTEIN (FRAGMENT).  
 GN SC2H4.01.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID:1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Oliver K., Harris D.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE:97000351; PubMed-8843436;  
 RA Redenbach M., Kleser H.M., Denapite D., Etchiner A., Cullum J.,  
 RA Kinashil H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL031514; CA20595.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 1  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SO SEQUENCE 606 AA; 63662 MW; 42ADF496A9F39CCD CRC64;

Query Match 54.58; Score 6; DB 2; Length 606;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKAP 9  
 |||||  
 DB 4 TAPKAP 9

RESULT 16  
 ID 092614 PRELIMINARY; PRT; 612 AA.  
 AC 092614;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE DIHYDROLIPOAMIDE ACETYLTRANSFERASE.  
 GN PDHB.  
 OS Streptomyces scouleri.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID:73044;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE:99078078; PubMed-9858775;  
 RA Youn H., Kwak J., Youn H.D., Hah Y.C., Kang S.O.;  
 RT "Lipoamide dehydrogenase from Streptomyces scouleri: biochemical and  
 genetic properties.";  
 RL Biochim. Biophys. Acta 1388:405-418(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Youn H., Kang S.-O.;  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS LIPOYL-BINDING DOMAIN(S).  
 DR EMBL: AF047034; AAD17484.2; -;  
 DR HSSP: P11961; IEBD.  
 DR InterPro: IPR000089; -;  
 DR InterPro: IPR000531; -;  
 DR InterPro: IPR001078; -;  
 DR InterPro: IPR001978; -;  
 DR InterPro: IPR002965; -;  
 DR InterPro: IPR003016; -;  
 DR Pfam: PF00198; 2-oxoacid\_dh; 1.  
 DR Pfam: PF00364; biotin\_lipoyl; 2.  
 DR PRINTS: PRO1217; PRICHEXTENSN.  
 DR PRODOM: PD001115; -; 1.  
 DR PROSITE: PS00189; LIPOYL; 2.  
 DR PROSITE: PS00430; TONB\_DEPENDENT\_REC\_1; UNKNOWN.1.  
 KW Lipoyl; Transferase.  
 SO SEQUENCE 612 AA; 61367 MW; B804109089943056 CRC64;

Query Match 54.58; Score 6; DB 2; Length 612;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOQTAP 6  
 |||||  
 DB 262 QOQTAP 267

```

RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-94150718; PubMed-7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten T., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mieg J., Thomas K., Vaughan M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Minx P., Le T.;
RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: U50197; AAA91256.1; -.
DR HSSP: P23911; IC80.
SQ SEQUENCE 343 AA; 38845 MW; 239065057C69E8A0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 5; Length 343;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PKAPTE 11
|||||
DB 4 PKAPTE 9

RESULT 12
O9FPA4 PRELIMINARY; PRT; 348 AA.
AC O9FPA4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PUTATIVE HEAT SHOCK FACTOR PROTEIN 1 (HSF 1).
GN P0006C01.9.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
OC Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPPONBARE;
RA Saeki T., Matsunoto T., Yamamoto K.;
RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0006C01.";
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AP002744; BAB19067.1; -.
KW Heat shock.
SQ SEQUENCE 348 AA; 37756 MW; 98C61C272160A134 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 348;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOQTAP 6
|||||
DB 21 QOQTAP 26

RESULT 13
O20966 PRELIMINARY; PRT; 371 AA.
O20966

```

```

AC Q20966;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE F58B3.7 PROTEIN.
GN F58B3.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B.;
RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-94150718; PubMed-7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten T., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mieg J., Thomas K., Vaughan M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL: Z73427; CAAG7799.1; -.
DR InterPro: IPR000467; -.
DR InterPro: IPR000504; -.
DR Pfam: PR00076; rrm; 1.
DR Pfam: PF01585; G-patch; 1.
DR SMART: SM00360; RRM; 1.
SQ SEQUENCE 371 AA; 41207 MW; 7789A85A10CCCEC3D CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 5; Length 371;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAPT 10
|||||
DB 255 APKAPT 260

RESULT 14
O9VMV9 PRELIMINARY; PRT; 388 AA.
AC O9VMV9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CG18623 PROTEIN.
GN CG18623.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

```



Query Match 90.9%; Score 10; DB 2; Length 84;  
 Best Local Similarity 100.0%; Pred. No. 0.00014;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OOTAPKAPTE 11  
 |||||  
 DB 1 OOTAPKAPTE 10

## RESULT 4

O48775 PRELIMINARY; PRT; 86 AA.

AC O48775; 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 GN IAP.  
 OS Listeria monocytogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-12705 /89, TYPE 1;  
 RX MEDLINE=96118685; PubMed=7496516;  
 RA Rasmussen O.F., Skoldboe P., Dons L., Rossen L., Olsen J.E.;  
 RT "Listeria monocytogenes exists in at least three evolutionary lines:  
 evidence from flagellin, invasive associated protein and listeriolysin  
 O genes.";  
 RT O genes.";  
 RL Microbiology 141:2053-2061(1995).  
 DR EMBL: X85859; CAA59883.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 86  
 SQ SEQUENCE 86 AA; 8842 MW; AEE753BE7EA6D6F0 CRC64;

Query Match 90.9%; Score 10; DB 2; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 0.00014;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OOTAPKAPTE 11  
 |||||  
 DB 1 OOTAPKAPTE 10

## RESULT 5

O48782 PRELIMINARY; PRT; 86 AA.

AC O48782; 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 GN IAP.  
 OS Listeria monocytogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-7785, TYPE 2;  
 RX MEDLINE=96118685; PubMed=7496516;  
 RA Rasmussen O.F., Skoldboe P., Dons L., Rossen L., Olsen J.E.;  
 RT "Listeria monocytogenes exists in at least three evolutionary lines:  
 evidence from flagellin, invasive associated protein and listeriolysin  
 O genes.";  
 RT O genes.";  
 RL Microbiology 141:2053-2061(1995).  
 DR EMBL: X85877; CAA59889.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 86  
 SQ SEQUENCE 86 AA; 8784 MW; DBA33D02B0E797B7 CRC64;

Query Match 90.9%; Score 10; DB 2; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 0.00014;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OOTAPKAPTE 11  
 |||||  
 DB 1 OOTAPKAPTE 10

## RESULT 6

O48781 PRELIMINARY; PRT; 88 AA.

AC O48781; 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 GN IAP.  
 OS Listeria monocytogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-64.1.1, TYPE 2;  
 RX MEDLINE=96118685; PubMed=7496516;  
 RA Rasmussen O.F., Skoldboe P., Dons L., Rossen L., Olsen J.E.;  
 RT "Listeria monocytogenes exists in at least three evolutionary lines:  
 evidence from flagellin, invasive associated protein and listeriolysin  
 O genes.";  
 RT O genes.";  
 RL Microbiology 141:2053-2061(1995).  
 DR EMBL: X85875; CAA59887.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 88  
 SQ SEQUENCE 88 AA; 8999 MW; B61ACB649E465739 CRC64;

Query Match 90.9%; Score 10; DB 2; Length 88;  
 Best Local Similarity 100.0%; Pred. No. 0.00014;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OOTAPKAPTE 11  
 |||||  
 DB 1 OOTAPKAPTE 10

## RESULT 7

O48774 PRELIMINARY; PRT; 117 AA.

AC O48774; 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 GN IAP.  
 OS Listeria monocytogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Listeria.  
 OX NCBI\_TaxID=1639;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-12067, TYPE 1;  
 RX MEDLINE=96118685; PubMed=7496516;  
 RA Rasmussen O.F., Skoldboe P., Dons L., Rossen L., Olsen J.E.;  
 RT "Listeria monocytogenes exists in at least three evolutionary lines:  
 evidence from flagellin, invasive associated protein and listeriolysin  
 O genes.";  
 RT O genes.";  
 RL Microbiology 141:2053-2061(1995).  
 DR EMBL: X85856; CAA59870.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 117

```
969 4 36.4 137 14 036896 036896 human immun
970 4 36.4 137 14 065517 065517 barmah fore
971 4 36.4 137 14 093671 093671 human immun
972 4 36.4 137 14 09E2J0 09E2J0 hepatitis c
973 4 36.4 137 14 09E2I9 09E2I9 hepatitis c
974 4 36.4 138 3 094376 094376 schizosacch
975 4 36.4 138 4 09NV55 09NV55 homo sapien
976 4 36.4 138 5 09VVP6 09VVP6 drosophila
977 4 36.4 138 10 043217 043217 triticum ae
978 4 36.4 138 10 09ZUS0 09ZUS0 arabidopsis
979 4 36.4 138 10 09LEF6 09LEF6 arabidopsis
980 4 36.4 138 14 09YYH4 09YYH4 human immun
981 4 36.4 139 14 09WS65 09WS65 human immun
982 4 36.4 139 2 09KZ29 09KZ29 streptomyc
983 4 36.4 139 2 09FA78 09FA78 vibrio pela
984 4 36.4 139 2 09FA76 09FA76 vibrio orle
985 4 36.4 139 3 09USJ1 09USJ1 schizosacch
986 4 36.4 139 5 09VM28 09VM28 drosophila
987 4 36.4 139 5 09U6U3 09U6U3 plasmodium
988 4 36.4 139 5 09NJM3 09NJM3 eumesocampa
989 4 36.4 139 5 09N1L1 09N1L1 machiloides
990 4 36.4 139 5 09N1J3 09N1J3 speleonecte
991 4 36.4 139 5 09N1J1 09N1J1 lomocerus s
992 4 36.4 139 5 09GQ22 09GQ22 plasmodium
993 4 36.4 139 10 065818 065818 lycopersico
994 4 36.4 139 10 09LL86 09LL86 solanum tub
995 4 36.4 139 10 09FK03 09FK03 arabidopsis
996 4 36.4 140 5 030082 030082 archaeoglob
997 4 36.4 140 5 025703 025703 plasmodium
998 4 36.4 140 5 018318 018318 chiromomus
999 4 36.4 140 10 065052 065052 picea maria
1000 4 36.4 140 10 09XETO 09XETO glycine max
```

## ALIGNMENTS

```
RESULT 1
ID 003493 PRELIMINARY; PRT; 478 AA.
AC 003493;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PROTEIN P60 PRECURSOR (INVASION-ASSOCIATED PROTEIN).
GN IAP.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP MEDLINE=93094153; PubMed=1459966;
RA Bubert A., Kuhn M., Goebel W., Koehler S.;
RT "Structural and functional properties of the p60 proteins from
RT different Listeria species."
RL J. Bacteriol. 174:8166-8171(1992).
CC -1- FUNCTION: THIS MAJOR EXTRACELLULAR PROTEIN MAY BE INVOLVED IN THE
CC INVASION OF NONPROFESSIONAL PHAGOCYTTIC CELLS BY LISTERIA.
CC EMBL: M80351; AAA25280.1; -.
DR InterPro: IPR000064; -.
DR InterPro: IPR002482; -.
DR Pfam: PF00877; NRPC_P60; 1.
DR Pfam: PF01476; LysM; 2.
DR SMART: SM00257; LysM; 1.
KM Signal.
FT SIGNAL. 1 27 BY SIMILARITY.
FT CHAIN 28 478 PROTEIN P60.
FT SEQUENCE 478 AA; 49997 MW; D46D050507402344 CRC64;
```

Query Match 100.0%; Score 11; DB 2; Length 478;  
Best Local Similarity 100.0%; Pred. No. 4.2e-05;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 00OTAPKAPTE 11

Db 291 00OTAPKAPTE 301

## RESULT 2

```
048784
ID 048784 PRELIMINARY; PRT; 76 AA.
AC 048784;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE INVASIVE ASSOCIATED PROTEIN (FRAGMENT).
GN IAP.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=443, TYPE 3;
RA Rasmussen O.F., Skouboe P., Dons L., Rossen L., Olsen J.E.;
RT "Listeria monocytogenes exists in at least three evolutionary lines:
RT evidence from flagellin, invasive associated protein and listeriolysin
RT O genes."
RL MicrobiolDgy 141:2053-2061(1995).
DR EMBL: X85880; CAA59892.1; -.
FT NON_TER 1
FT NON_TER 76
FT SEQUENCE 76 AA; 7744 MW; D5AC5B516193652E CRC64;
```

Query Match 90.9%; Score 10; DB 2; Length 76;  
Best Local Similarity 100.0%; Pred. No. 0.00013;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 00TAPKAPTE 11

Db 1 00TAPKAPTE 10

## RESULT 3

```
057205
ID 057205 PRELIMINARY; PRT; 84 AA.
AC 057205;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE INVASIVE ASSOCIATED PROTEIN (FRAGMENT).
GN IAP.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=10165, TYPE 1;
RA Rasmussen O.F., Skouboe P., Dons L., Rossen L., Olsen J.E.;
RT "Listeria monocytogenes exists in at least three evolutionary lines:
RT evidence from flagellin, invasive associated protein and listeriolysin
RT O genes."
RL Microbiology 141:2053-2061(1995).
DR EMBL: X85859; CAA59873.1; -.
DR EMBL: X85857; CAA59871.1; -.
DR EMBL: X85858; CAA59872.1; -.
FT NON_TER 1
FT NON_TER 84
FT SEQUENCE 84 AA; 8627 MW; 8796F59721315E3C CRC64;
```

823	4	36.4	119	5	09N6G3	09n6g3 trypanosoma	896	4	36.4	128	5	09NIO2	09niq2 trypanosoma
824	4	36.4	119	10	039332	039332 brassica na	897	4	36.4	128	5	09N2V3	09n2v3 caenorhabdi
825	4	36.4	119	10	039404	039404 brassica ra	898	4	36.4	128	10	09FY37	09fy37 brassica ol
826	4	36.4	119	10	09SJW3	09sjw3 arabadopsis	899	4	36.4	128	13	P88190	p88190 carcharinu
827	4	36.4	119	10	09STR1	09str1 arabadopsis	900	4	36.4	128	14	09JG52	09jg52 tlw-like ml
828	4	36.4	119	10	09FJUI	09fjui arabadopsis	901	4	36.4	129	2	057111	057111 veillonella
829	4	36.4	119	13	09W7A3	09w7a3 coturnix co	902	4	36.4	129	5	09Y1V5	09y1v5 ephydactia f
830	4	36.4	119	13	09PWB2	09pwb2 coturnix co	903	4	36.4	129	5	09V1R5	09v1r5 ephydactia f
831	4	36.4	119	14	037047	037047 human immun	904	4	36.4	129	14	036883	036883 human immun
832	4	36.4	119	14	076134	076134 human immun	905	4	36.4	130	1	09Y9J3	09y9j3 pyrococcus
833	4	36.4	119	14	09J6R6	09j6r6 human immun	906	4	36.4	130	2	047852	047852 erwinia her
834	4	36.4	119	14	09J6R2	09j6r2 human immun	907	4	36.4	130	2	045333	045333 brucella ov
835	4	36.4	120	2	P76515	P76515 escherichia	908	4	36.4	130	2	056729	056729 vibrio sp.
836	4	36.4	120	2	09ZNI2	09zni2 staphylococ	909	4	36.4	130	2	066857	066857 aquifex ae
837	4	36.4	120	5	09NIO7	09niq7 trypanosoma	910	4	36.4	130	2	09S347	09s347 prevotella
838	4	36.4	120	6	077637	077637 sus scrofa	911	4	36.4	130	5	03D410	03d410 manduca sex
839	4	36.4	120	10	09LJW6	09ljw6 arabadopsis	912	4	36.4	130	6	09TSY4	09tsy4 sus scrofa
840	4	36.4	120	10	09J019	09j019 human immun	913	4	36.4	130	9	092XJ0	092xj0 bacterioph
841	4	36.4	120	14	09J018	09j018 human immun	914	4	36.4	130	13	09PU29	09pu29 struthio ca
842	4	36.4	121	2	084349	084349 chlamydia t	915	4	36.4	131	2	09RSV2	09rsv2 delinococcus
843	4	36.4	121	2	P74135	P74135 synechocyst	916	4	36.4	131	2	085701	085701 streptomyce
844	4	36.4	121	2	09PRD4	09prd4 ureaplasma	917	4	36.4	131	4	09S331	09s331 homo sapien
845	4	36.4	121	5	09W306	09w306 drosophila	918	4	36.4	131	5	09N131	09n131 plasmodium
846	4	36.4	121	5	09W2S2	09w2s2 drosophila	919	4	36.4	132	2	052174	052174 salmonella
847	4	36.4	121	6	09TSD5	09tsd5 bos taurus	920	4	36.4	132	3	09P323	09p323 pneumocysti
848	4	36.4	121	8	034194	034194 crithidia o	921	4	36.4	132	4	016609	016609 homo sapien
849	4	36.4	121	10	041575	041575 tritlicum ae	922	4	36.4	132	5	09NIO3	09niq3 trypanosoma
850	4	36.4	121	10	09M573	09m573 perilla fru	923	4	36.4	132	5	09N803	09n803 leishmania
851	4	36.4	121	13	09PWB4	09pwb4 coturnix co	924	4	36.4	132	5	09GR02	09gr02 plasmodium
852	4	36.4	121	13	09PWB1	09pwb1 coturnix co	925	4	36.4	132	10	09LZ45	09l245 arabadopsis
853	4	36.4	121	14	072893	072893 human immun	926	4	36.4	132	11	060581	060581 mus musculu
854	4	36.4	121	14	09WB17	09wb17 human immun	927	4	36.4	132	14	036856	036856 human immun
855	4	36.4	122	2	059657	059657 methanosarc	928	4	36.4	132	14	097693	097693 human immun
856	4	36.4	122	2	09S173	09s173 streptococc	929	4	36.4	132	14	09YFD3	09yfd3 human immun
857	4	36.4	122	2	09RA82	09ra82 streptococc	930	4	36.4	132	14	09WL78	09wl78 human immun
858	4	36.4	122	2	09RA81	09ra81 streptococc	931	4	36.4	132	14	09WL77	09wl77 human immun
859	4	36.4	122	2	09R391	09r391 streptococc	932	4	36.4	132	14	09JH74	09jht4 chimpanzee
860	4	36.4	122	2	09R350	09r350 streptococc	933	4	36.4	133	3	005666	005666 saccharomyc
861	4	36.4	122	5	097361	097361 phytophthor	934	4	36.4	133	14	09YRE8	09yre8 human immun
862	4	36.4	122	5	09G6W4	09g6w4 leishmania	935	4	36.4	133	14	09YTA7	09yta7 human immun
863	4	36.4	122	13	09DFN7	09dfn7 gilllichthys	936	4	36.4	134	1	09YFO6	09yfo6 aetopyrum p
864	4	36.4	123	1	09YB08	09ybu8 aetopyrum p	937	4	36.4	134	2	09YKLA	09ykl8 vibrio chol
865	4	36.4	123	2	024882	024882 helicobacte	938	4	36.4	134	2	09JZ33	09jz33 neisseria m
866	4	36.4	123	2	09RXU9	09rxu9 delinococcus	939	4	36.4	134	4	09NV29	09nv29 homo sapien
867	4	36.4	123	4	09H7H2	09h7h2 homo sapien	940	4	36.4	134	5	024959	024959 giardia lam
868	4	36.4	123	5	09VE70	09ve70 drosophila	941	4	36.4	134	5	024988	024988 giardia lam
869	4	36.4	124	2	09R2F1	09r2f1 bruceia ab	942	4	36.4	134	5	09GX38	09gx38 leishmania
870	4	36.4	124	4	09H1H0	09h1h0 homo sapien	943	4	36.4	134	6	09N2E7	09n2e7 gorilla gor
871	4	36.4	124	5	097359	097359 phytophthor	944	4	36.4	134	10	09FSP1	09fsp1 oryza sativ
872	4	36.4	124	9	038372	038372 lactococcus	945	4	36.4	134	11	070483	070483 rattus norv
873	4	36.4	124	11	062286	062286 mus musculu	946	4	36.4	134	14	097679	097679 human immun
874	4	36.4	124	13	09DFD3	09dfd3 oncothrychnu	947	4	36.4	135	2	053977	053977 mycobacteri
875	4	36.4	124	13	09J4Z3	09j4z3 fowlpox vir	948	4	36.4	135	4	09P207	09p2q7 homo sapien
876	4	36.4	125	1	09Y973	09y973 aetopyrum p	949	4	36.4	135	4	09H524	09h524 homo sapien
877	4	36.4	125	1	P72987	P72987 synechocyst	950	4	36.4	135	6	09N2E8	09n2e8 pan troglod
878	4	36.4	125	10	09FRS9	09frs9 arabadopsis	951	4	36.4	135	6	09N2E6	09n2e6 pongo pygma
879	4	36.4	125	13	09W7A4	09w7a4 coturnix co	952	4	36.4	135	10	043215	043215 tritlicum ae
880	4	36.4	125	13	09PWB5	09pwb5 coturnix co	953	4	36.4	135	10	09YVE0	09yve0 arabadopsis
881	4	36.4	125	13	09PWB3	09pwb3 coturnix co	954	4	36.4	135	13	09W7A6	09w7a6 coturnix co
882	4	36.4	125	14	09WB44	09wb44 human immun	955	4	36.4	136	1	09YAR0	09yar0 aetopyrum p
883	4	36.4	125	14	069271	069271 equine herp	956	4	36.4	136	2	051792	051792 alcaligenes
884	4	36.4	125	14	09DOW8	09dow8 human immun	957	4	36.4	136	2	P74444	P74444 synechocyst
885	4	36.4	126	2	P74205	P74205 synechocyst	958	4	36.4	136	5	09G0Z8	09g0z8 plasmodium
886	4	36.4	126	4	09UHE2	09uhe2 homo sapien	959	4	36.4	136	10	093318	093318 brassica na
887	4	36.4	126	4	094720	094720 paramedcum	960	4	36.4	136	10	093319	093319 brassica na
888	4	36.4	126	10	09SEF5	09sef5 arabadopsis	961	4	36.4	136	10	043216	043216 tritlicum ae
889	4	36.4	127	2	09RLC9	09rlc9 pseudomonas	962	4	36.4	136	10	09S7N4	09s7n4 arabadopsis
890	4	36.4	127	2	09KLS8	09kls8 vibrio chol	963	4	36.4	136	10	09FX60	09fx60 arabadopsis
891	4	36.4	127	5	09N128	09n128 plasmodium	964	4	36.4	137	2	047836	047836 enterococcu
892	4	36.4	127	10	09ZU00	09zu00 arabadopsis	965	4	36.4	137	5	015689	015689 plasmodium
893	4	36.4	127	13	09W7A5	09w7a5 coturnix co	966	4	36.4	137	5	020468	020468 caenorhabdi
894	4	36.4	128	2	09WY58	09wy58 thermotoga	967	4	36.4	137	5	09NIO4	09niq4 trypanosoma
895	4	36.4	128	2	09KXJ3	09kxj3 streptomyc	968	4	36.4	137	11	035946	035946 rattus norv



677	4	36.4	102	2	053562	053562 methylolisin	750	4	36.4	112	14	079829	079829 human immun
678	4	36.4	102	5	09VED9	09VED9 drosophila	751	4	36.4	112	14	079836	079836 human immun
679	4	36.4	102	14	079847	079847 human immun	752	4	36.4	113	1	059655	059655 methanosarc
680	4	36.4	102	14	079848	079848 human immun	753	4	36.4	113	6	0913F9	0913F9 neptiroselmt
681	4	36.4	102	14	079849	079849 human immun	754	4	36.4	113	14	071164	071164 hepatitis c
682	4	36.4	102	14	091E50	091E50 human immun	755	4	36.4	113	14	079852	079852 human immun
683	4	36.4	103	1	059510	059510 pyrococcus	756	4	36.4	113	14	079853	079853 human immun
684	4	36.4	103	2	09PHX9	09PHX9 campylobact	757	4	36.4	113	14	079855	079855 human immun
685	4	36.4	103	2	09KRB8	09KRB8 vibrio chol	758	4	36.4	113	14	079856	079856 human immun
686	4	36.4	103	3	059673	059673 schizosacch	759	4	36.4	113	14	079857	079857 human immun
687	4	36.4	103	3	09E699	09E699 neurospora	760	4	36.4	113	14	079858	079858 human immun
688	4	36.4	103	5	022338	022338 caenorhabd1	761	4	36.4	113	14	079872	079872 human immun
689	4	36.4	103	5	09VZ98	09VZ98 drosophila	762	4	36.4	113	14	079873	079873 human immun
690	4	36.4	103	5	09NIO8	09NIO8 trypanosoma	763	4	36.4	113	14	079858	079858 human immun
691	4	36.4	103	5	09NIO8	09NIO8 trypanosoma	764	4	36.4	113	14	079859	079859 human immun
692	4	36.4	103	5	09NIO5	09NIO5 trypanosoma	765	4	36.4	113	14	079860	079860 human immun
693	4	36.4	104	3	09NIO2	09NIO2 plasmodium	766	4	36.4	113	14	079861	079861 human immun
694	4	36.4	104	8	09MED0	09MED0 beta vulgar	767	4	36.4	113	14	079862	079862 human immun
695	4	36.4	104	10	09SUD3	09SUD3 arabidopsis	768	4	36.4	113	14	079870	079870 human immun
696	4	36.4	105	2	09MWP7	09MWP7 synchococc	769	4	36.4	113	14	079871	079871 human immun
697	4	36.4	105	2	025027	025027 helicobacte	770	4	36.4	114	2	09RRC3	09RRC3 listeria mo
698	4	36.4	105	2	09JTK2	09JTK2 neisseria m	771	4	36.4	114	10	023138	023138 arabidopsis
699	4	36.4	105	5	09VRR8	09VRR8 drosophila	772	4	36.4	114	14	079854	079854 human immun
700	4	36.4	105	14	076624	076624 human immun	773	4	36.4	115	5	09N129	09N129 plasmodium
701	4	36.4	105	14	012332	012332 human immun	774	4	36.4	115	5	09N130	09N130 plasmodium
702	4	36.4	105	14	09YTT7	09YTT7 human immun	775	4	36.4	115	5	09N129	09N129 plasmodium
703	4	36.4	105	14	09QTV8	09QTV8 tt virus. o	776	4	36.4	115	10	09XU07	09XU07 arabidopsis
704	4	36.4	105	14	09Q1T3	09Q1T3 human immun	777	4	36.4	115	10	09LYA0	09LYA0 arabidopsis
705	4	36.4	105	14	09Q1T2	09Q1T2 human immun	778	4	36.4	115	10	09LTR0	09LTR0 arabidopsis
706	4	36.4	105	14	09PPW9	09PPW9 human immun	779	4	36.4	116	1	09YA88	09YA88 aeropyrum p
707	4	36.4	106	1	027467	027467 methanobact	780	4	36.4	116	3	09XAW0	09XAW0 pseudomonas
708	4	36.4	106	1	09YBC9	09YBC9 aeropyrum p	781	4	36.4	116	3	043034	043034 schizosacch
709	4	36.4	106	5	09Y9E1	09Y9E1 aeropyrum p	782	4	36.4	116	4	013138	013138 homo sapien
710	4	36.4	106	5	09NTO6	09NTO6 trypanosoma	783	4	36.4	116	5	P90590	P90590 parametium
711	4	36.4	106	10	065746	065746 ciccer arlet	784	4	36.4	116	5	09GQ27	09GQ27 plasmodium
712	4	36.4	106	10	09EY11	09EY11 picea abies	785	4	36.4	116	5	09GQ26	09GQ26 plasmodium
713	4	36.4	106	14	085590	085590 human t-cel	786	4	36.4	116	8	09XDF5	09XDF5 gonostoma g
714	4	36.4	106	14	098499	098499 parametium	787	4	36.4	116	10	09XFT6	09XFT6 chlamydomon
715	4	36.4	107	5	09M4Z4	09M4Z4 drosophila	788	4	36.4	116	10	09SNU1	09SNU1 cryza sativ
716	4	36.4	107	5	09WOT6	09WOT6 drosophila	789	4	36.4	116	14	079826	079826 human immun
717	4	36.4	107	5	09U393	09U393 caenorhabd1	790	4	36.4	116	14	079830	079830 human immun
718	4	36.4	107	5	09NLP7	09NLP7 leishmania	791	4	36.4	116	14	079832	079832 human immun
719	4	36.4	107	5	09NEC8	09NEC8 leishmania	792	4	36.4	116	14	079834	079834 human immun
720	4	36.4	107	5	09XXL3	09XXL3 caenorhabd1	793	4	36.4	116	14	079835	079835 human immun
721	4	36.4	107	10	09ZRX4	09ZRX4 tritlicum ae	794	4	36.4	116	14	09YYV1	09YYV1 eyach virus
722	4	36.4	108	8	09ZYO8	09ZYO8 chordelies	795	4	36.4	116	14	09YYV0	09YYV0 eyach virus
723	4	36.4	109	2	P73607	P73607 synchocyst	796	4	36.4	117	2	0910H1	0910H1 pseudomonas
724	4	36.4	109	4	014325	014325 homo sapien	797	4	36.4	117	4	09NTE4	09NTE4 homo sapien
725	4	36.4	109	10	042347	042347 arabidopsis	798	4	36.4	117	4	095166	095166 homo sapien
726	4	36.4	110	2	09PEW3	09PEW3 xylella fas	799	4	36.4	117	4	09H0R8	09H0R8 homo sapien
727	4	36.4	110	2	09PEW2	09PEW2 neurospora	800	4	36.4	117	5	09VAX4	09VAX4 drosophila
728	4	36.4	110	5	09N316	09N316 caenorhabd1	801	4	36.4	117	10	048094	048094 leavenworth
729	4	36.4	110	6	09GMW6	09GMW6 macaca fasc	802	4	36.4	117	10	050049	050049 leavenworth
730	4	36.4	110	10	081061	081061 arabidopsis	803	4	36.4	117	10	050059	050059 leavenworth
731	4	36.4	110	11	09Z172	09Z172 mus musculu	804	4	36.4	117	10	050060	050060 leavenworth
732	4	36.4	110	14	011343	011343 molluscum c	805	4	36.4	117	11	09QUIT7	09QUIT7 rattus norv
733	4	36.4	110	14	079840	079840 human immun	806	4	36.4	117	11	09J397	09J397 mus musculu
734	4	36.4	110	14	079841	079841 human immun	807	4	36.4	117	14	098304	098304 molluscum c
735	4	36.4	110	14	079842	079842 human immun	808	4	36.4	117	14	079831	079831 human immun
736	4	36.4	110	14	P89815	P89815 human immun	809	4	36.4	117	14	079833	079833 human immun
737	4	36.4	111	5	094666	094666 plasmodium	810	4	36.4	118	2	024992	024992 helicobacte
738	4	36.4	111	5	09N136	09N136 plasmodium	811	4	36.4	118	2	09ZWN3	09ZWN3 helicobacte
739	4	36.4	111	5	09N135	09N135 plasmodium	812	4	36.4	118	2	09S1B5	09S1B5 arthropacte
740	4	36.4	111	5	09N134	09N134 plasmodium	813	4	36.4	118	4	09YGB1	09YGB1 homo sapien
741	4	36.4	111	10	024263	024263 pisum sativ	814	4	36.4	118	5	017150	017150 branchiostor
742	4	36.4	111	10	09SPK9	09SPK9 lactuca sat	815	4	36.4	118	5	097360	097360 phytophthor
743	4	36.4	112	1	09HNA4	09HNA4 halobacteri	816	4	36.4	118	5	09N1P8	09N1P8 solanum tub
744	4	36.4	112	2	09R2M7	09R2M7 delnoccocc	817	4	36.4	118	10	049947	049947 arbidopsis
745	4	36.4	112	5	09JTH5	09JTH5 neisseria m	818	4	36.4	118	10	09ZRD4	09ZRD4 phytophthor
746	4	36.4	112	5	09GVD6	09GVD6 leishmania	819	4	36.4	118	10	09ROM7	09ROM7 arbidopsis
747	4	36.4	112	14	065509	065509 barmanah fore	820	4	36.4	118	10	09SMZ7	09SMZ7 arbidopsis
748	4	36.4	112	14	079827	079827 human immun	821	4	36.4	119	1	0919V6	0919V6 aeropyrum p
749	4	36.4	112	14	079828	079828 human immun	822	4	36.4	119	2	055617	055617 synchocyst

531	4	36.4	71	2	Q9PAE9	Q9paeg xyliella fas	604	4	36.4	84	14	Q9QJP8	Q9qjp8 human immun
532	4	36.4	71	10	Q9FTW6	Q9ftw6 oryza sativ	605	4	36.4	84	14	Q9QJP6	Q9qjp6 human immun
533	4	36.4	72	2	Q4J953	Q4j953 bacillus br	606	4	36.4	84	14	Q9QJP5	Q9qjp5 human immun
534	4	36.4	72	3	Q00411	Q00411 coccidioid	607	4	36.4	84	14	Q9QJP4	Q9qjp4 human immun
535	4	36.4	72	7	Q9EP71	Q9ep71 gadus morhu	608	4	36.4	84	14	Q9QJP3	Q9qjp3 human immun
536	4	36.4	72	14	Q84416	Q84416 paramecium	609	4	36.4	84	14	Q9QJP2	Q9qjp2 human immun
537	4	36.4	73	4	Q95576	Q95576 homo sapien	610	4	36.4	84	14	Q9QJP1	Q9qjp1 human immun
538	4	36.4	73	10	Q9M9R9	Q9m9r9 arabidopsis	611	4	36.4	84	14	Q9QJN9	Q9qjn9 human immun
539	4	36.4	74	2	Q43946	Q43946 azorhizobiu	612	4	36.4	84	14	Q9QJN8	Q9qjn8 human immun
540	4	36.4	74	2	Q31653	Q31653 bacillus su	613	4	36.4	85	6	Q29356	Q29356 sus scrofa
541	4	36.4	74	2	Q32280	Q32280 bacillus su	614	4	36.4	85	9	Q9WC17	Q9wc17 bacterioph
542	4	36.4	75	2	Q44890	Q44890 borellia bu	615	4	36.4	86	4	Q99557	Q99557 homo sapien
543	4	36.4	75	6	P79319	P79319 sus scrofa	616	4	36.4	86	8	Q20119	Q20119 chlorella v
544	4	36.4	76	2	Q06243	Q06243 mycobacteri	617	4	36.4	86	14	Q00792	Q00792 human immun
545	4	36.4	76	6	Q9N046	Q9n046 macaca fasc	618	4	36.4	87	5	Q9U3U6	Q9u3u6 asterias ru
546	4	36.4	76	10	P93382	P93382 nicotiana t	619	4	36.4	87	14	Q65930	Q65930 caprine art
547	4	36.4	77	2	Q9X6M2	Q9x6m2 klebsiella	620	4	36.4	87	14	Q9WR53	Q9wr53 brazilian c
548	4	36.4	77	2	Q9RVN7	Q9rvn7 delinococcus	621	4	36.4	87	14	Q9WR51	Q9wr51 brazilian c
549	4	36.4	77	10	Q9FMI6	Q9fmi6 arabidopsis	622	4	36.4	87	14	Q9WR49	Q9wr49 brazilian c
550	4	36.4	78	10	Q9LDR0	Q9ldr0 oryza sativ	623	4	36.4	87	14	Q9W959	Q9w959 brazilian c
551	4	36.4	78	14	Q79844	Q79844 human immun	624	4	36.4	88	2	Q56971	Q56971 yersinia pe
552	4	36.4	78	14	Q79846	Q79846 human immun	625	4	36.4	88	2	Q9XA88	Q9xa88 streptomyc
553	4	36.4	79	5	Q23145	Q23145 caenorhabdi	626	4	36.4	88	5	Q9VFC8	Q9vfc8 drosophila
554	4	36.4	79	11	Q63280	Q63280 rattus norv	627	4	36.4	89	1	Q9YAL2	Q9yal2 aeropyrum p
555	4	36.4	79	14	Q9QJ58	Q9qj58 human herpe	628	4	36.4	89	1	Q9HSU7	Q9hsu7 halobacteri
556	4	36.4	81	5	Q22861	Q22861 caenorhabdi	629	4	36.4	89	13	Q42371	Q42371 brachydanio
557	4	36.4	82	10	Q9SY18	Q9sy18 arabidopsis	630	4	36.4	89	14	Q79821	Q79821 human immun
558	4	36.4	82	14	Q55446	Q55446 sindbis vir	631	4	36.4	90	5	Q97102	Q97102 drosophila
559	4	36.4	82	14	Q79824	Q79824 human immun	632	4	36.4	90	14	Q79819	Q79819 human immun
560	4	36.4	82	14	Q79623	Q79623 human immun	633	4	36.4	90	14	Q92982	Q92982 chlamydia p
561	4	36.4	83	14	Q79825	Q79825 human immun	634	4	36.4	91	2	Q9RZ89	Q9rzr9 delinococcus
562	4	36.4	83	1	Q9NHYS	Q9nhys halobacteri	635	4	36.4	91	2	Q9R6C6	Q9r6c6 acrobacteri
563	4	36.4	83	1	Q9NHJ1	Q9nhj1 halobacteri	636	4	36.4	91	2	Q9JST1	Q9jst1 chlamydia p
564	4	36.4	83	2	Q85409	Q85409 pseudomonas	637	4	36.4	91	14	Q69372	Q69372 cercopithec
565	4	36.4	83	2	Q85410	Q85410 pseudomonas	638	4	36.4	91	14	Q9YMG6	Q9ymg6 lymantria d
566	4	36.4	83	2	Q85411	Q85411 pseudomonas	639	4	36.4	92	6	Q9TV11	Q9tv11 canis famli
567	4	36.4	83	2	Q85412	Q85412 pseudomonas	640	4	36.4	92	13	Q9W7A7	Q9w7a7 coturnix co
568	4	36.4	83	2	Q85413	Q85413 pseudomonas	641	4	36.4	93	5	Q9W2X9	Q9w2x9 schizosacch
569	4	36.4	83	2	Q85414	Q85414 pseudomonas	642	4	36.4	93	5	Q9TVV9	Q9tvv9 caenorhabdi
570	4	36.4	83	2	Q85415	Q85415 pseudomonas	643	4	36.4	93	5	Q9W2X9	Q9w2x9 drosophila
571	4	36.4	83	2	Q85416	Q85416 pseudomonas	644	4	36.4	93	5	Q9TVV9	Q9tvv9 caenorhabdi
572	4	36.4	83	2	Q85418	Q85418 pseudomonas	645	4	36.4	93	10	Q42409	Q42409 triticum tu
573	4	36.4	83	2	Q85419	Q85419 pseudomonas	646	4	36.4	93	10	P93609	P93609 triticum ae
574	4	36.4	83	2	Q85420	Q85420 pseudomonas	647	4	36.4	93	10	Q9SDN8	Q9sdsn phytophthor
575	4	36.4	83	2	Q85421	Q85421 pseudomonas	648	4	36.4	94	2	Q9RSY3	Q9rsy3 azorhizobiu
576	4	36.4	83	2	Q85422	Q85422 pseudomonas	649	4	36.4	94	10	Q40830	Q40830 pisum etati
577	4	36.4	83	2	Q85423	Q85423 pseudomonas	650	4	36.4	94	10	Q24255	Q24255 pisum fulvu
578	4	36.4	83	2	Q85424	Q85424 pseudomonas	651	4	36.4	94	10	Q24262	Q24262 pisum sativ
579	4	36.4	83	2	Q85425	Q85425 pseudomonas	652	4	36.4	94	10	Q9M4X1	Q9m4x1 poa fendler
580	4	36.4	83	2	Q85426	Q85426 pseudomonas	653	4	36.4	94	10	Q9FSJ9	Q9fsj9 oryza sativ
581	4	36.4	83	2	Q85427	Q85427 pseudomonas	654	4	36.4	95	2	Q32887	Q32887 mycobacteri
582	4	36.4	83	2	Q85428	Q85428 pseudomonas	655	4	36.4	95	14	Q86553	Q86553 maedi-vlsna
583	4	36.4	83	2	Q85429	Q85429 pseudomonas	656	4	36.4	95	14	Q79868	Q79868 human immun
584	4	36.4	83	2	Q85430	Q85430 pseudomonas	657	4	36.4	95	14	Q79869	Q79869 human immun
585	4	36.4	83	2	Q85431	Q85431 pseudomonas	658	4	36.4	96	5	Q9V5Y7	Q9v5y7 drosophila
586	4	36.4	83	2	Q85432	Q85432 pseudomonas	659	4	36.4	96	5	Q9VORT	Q9vort coturnix co
587	4	36.4	83	2	Q85433	Q85433 pseudomonas	660	4	36.4	97	13	Q9W7A8	Q9w7a8 coturnix co
588	4	36.4	83	2	Q85434	Q85434 pseudomonas	661	4	36.4	98	5	Q96731	Q96731 phytophthor
589	4	36.4	83	2	Q85435	Q85435 pseudomonas	662	4	36.4	98	5	Q96732	Q96732 phytophthor
590	4	36.4	83	2	Q85436	Q85436 pseudomonas	663	4	36.4	98	5	Q96733	Q96733 phytophthor
591	4	36.4	83	2	Q85437	Q85437 pseudomonas	664	4	36.4	98	5	Q96734	Q96734 phytophthor
592	4	36.4	83	2	Q85438	Q85438 pseudomonas	665	4	36.4	98	10	Q9S9F8	Q9s9f8 phytophthor
593	4	36.4	83	2	Q85439	Q85439 pseudomonas	666	4	36.4	98	14	Q57056	Q57056 human immun
594	4	36.4	83	2	Q85440	Q85440 pseudomonas	667	4	36.4	98	14	Q9YPO0	Q9ypo0 choristoneu
595	4	36.4	83	2	Q85441	Q85441 pseudomonas	668	4	36.4	99	2	Q92718	Q92718 chlamydia p
596	4	36.4	83	2	Q85442	Q85442 pseudomonas	669	4	36.4	99	2	Q9ERT5	Q9ert5 corynebacte
597	4	36.4	83	2	Q85443	Q85443 pseudomonas	670	4	36.4	99	3	Q9HPI1	Q9hpi1 picchia etch
598	4	36.4	83	2	Q85444	Q85444 pseudomonas	671	4	36.4	101	10	Q9W581	Q9w581 splinacia ol
599	4	36.4	83	2	Q85445	Q85445 pseudomonas	672	4	36.4	101	11	P70530	P70530 rattus norv
600	4	36.4	83	2	P77431	P77431 escherichia	673	4	36.4	101	11	Q9ERN9	Q9ern9 cavia porco
601	4	36.4	83	2	Q9KJ64	Q9kj64 vibrio chol	674	4	36.4	101	13	Q9PWB6	Q9pwb6 coturnix co
602	4	36.4	84	5	Q9XVZ8	Q9xvz8 caenorhabdi	675	4	36.4	101	13		
603	4	36.4	84	10	Q9XCQ9	Q9xcq9 oryza sativ	676	4	36.4	101	13		

385	5	45.5	1550	4	092547	092547 homo sapien	458	4	36.4	46	2	086393	086393 nelisseria m
386	5	45.5	1586	13	09DFB7	09dfb7 gallus gall	459	4	36.4	47	5	09V5D6	09v5d6 drosophila
387	5	45.5	1621	2	09KXPAS	09kxpas vibrio chol	460	4	36.4	47	14	087177	087177 chimpanzee
388	5	45.5	1654	11	09JIG0	09jig0 rattus norv	461	4	36.4	47	14	087178	087178 chimpanzee
389	5	45.5	1655	5	09VB5	09vb5 drosophila	462	4	36.4	47	14	087179	087179 chimpanzee
390	5	45.5	1663	5	09U001	09u001 homo sapien	463	4	36.4	47	14	087180	087180 chimpanzee
391	5	45.5	1664	5	09V9V5	09v9v5 drosophila	464	4	36.4	47	14	087181	087181 chimpanzee
392	5	45.5	1668	5	09N926	09n926 drosophila	465	4	36.4	47	14	087183	087183 chimpanzee
393	5	45.5	1696	11	09WTR8	09wtr8 rattus norv	466	4	36.4	47	14	087186	087186 chimpanzee
394	5	45.5	1703	5	09VGP8	09vgp8 drosophila	467	4	36.4	47	14	087188	087188 chimpanzee
395	5	45.5	1729	5	09U617	09u617 drosophila	468	4	36.4	48	2	09KX20	09kx20 pseudomonas
396	5	45.5	1785	5	02S685	02s685 plasmodium	469	4	36.4	48	4	012578	012578 homo sapien
397	5	45.5	1785	5	093636	093636 caenorhabdi	470	4	36.4	50	14	09YUW5	09yuw5 human coxa
398	5	45.5	1794	9	09TIA7	09tia7 bacterioph	471	4	36.4	50	14	09YUW4	09yuw4 echovirus 3
399	5	45.5	1825	3	09HFI5	09hfi5 neurospora	472	4	36.4	50	14	09YUW2	09yuw2 human coxa
400	5	45.5	1876	5	09VTR5	09vtr5 drosophila	473	4	36.4	50	14	09YUW0	09yuw0 human coxa
401	5	45.5	1929	5	093537	093537 caenorhabdi	474	4	36.4	52	14	09TMS5	09tms5 prunus necr
402	5	45.5	1990	5	09U800	09u800 drosophila	475	4	36.4	52	14	09TMS4	09tms4 prunus necr
403	5	45.5	1991	5	09W244	09w244 drosophila	476	4	36.4	52	14	09TMS2	09tms2 prunus necr
404	5	45.5	1994	10	09LP19	09lp19 arabidopsis	477	4	36.4	52	14	09TMS1	09tms1 prunus necr
405	5	45.5	2023	4	075557	075557 homo sapien	478	4	36.4	52	14	09TMR9	09tmr9 prunus necr
406	5	45.5	2033	4	09UMD7	09umd7 homo sapien	479	4	36.4	52	14	09TMR8	09tmr8 prunus necr
407	5	45.5	2074	11	088542	088542 mus musculu	480	4	36.4	52	14	09TMR6	09tmr6 prunus necr
408	5	45.5	2135	4	043157	043157 homo sapien	481	4	36.4	52	14	09TMR4	09tmr4 prunus necr
409	5	45.5	2135	4	09U1V7	09u1v7 homo sapien	482	4	36.4	52	14	09TMR3	09tmr3 prunus necr
410	5	45.5	2187	11	P70670	P70670 mus musculu	483	4	36.4	52	14	09TMR1	09tmr1 prunus necr
411	5	45.5	2209	5	09U0G6	09u0g6 plasmodium	484	4	36.4	52	14	09TMR0	09tmr0 prunus necr
412	5	45.5	2212	5	09UHV6	09uhv6 homo sapien	485	4	36.4	52	14	09TMR8	09tmr8 prunus necr
413	5	45.5	2212	5	094657	094657 plasmodium	486	4	36.4	52	14	09TMR6	09tmr6 prunus necr
414	5	45.5	2376	11	09JIF2	09jif2 rattus norv	487	4	36.4	52	14	09TMR5	09tmr5 prunus necr
415	5	45.5	2382	4	09H4A3	09h4a3 homo sapien	488	4	36.4	52	14	09TMR4	09tmr4 prunus necr
416	5	45.5	2439	5	09VMS2	09vms2 drosophila	489	4	36.4	52	14	09TMR4	09tmr4 prunus necr
417	5	45.5	2635	14	P88955	P88955 kaposi's sa	490	4	36.4	53	1	09H4M2	09h4m2 prunus necr
418	5	45.5	2635	14	040942	040942 kaposi's sa	491	4	36.4	53	1	09H4M2	09h4m2 prunus necr
419	5	45.5	2677	5	P90580	P90580 plasmodium	492	4	36.4	53	14	09H4M5	09h4m5 frankia sp.
420	5	45.5	2756	2	033904	033904 shewanella	493	4	36.4	55	6	097747	097747 bos taurus
421	5	45.5	2936	5	09NKP7	09nkp7 leishmania	494	4	36.4	55	6	097747	097747 bos taurus
422	5	45.5	2959	11	09JIF1	09jif1 rattus norv	495	4	36.4	56	3	09UVM9	09uvm9 coprinus ci
423	5	45.5	3112	5	09NKP1	09nkp1 leishmania	496	4	36.4	56	4	012976	012976 homo sapien
424	5	45.5	3132	14	P89459	P89459 herpes simp	497	4	36.4	57	6	097900	097900 ateles belz
425	5	45.5	3169	5	0917V8	0917v8 drosophila	498	4	36.4	59	4	012977	012977 homo sapien
426	5	45.5	3201	5	09NDJ2	09ndj2 drosophila	499	4	36.4	60	2	09PGC0	09pgc0 xylella fas
427	5	45.5	3257	5	09V736	09v736 drosophila	500	4	36.4	60	2	09KXN8	09kxn8 streptomyce
428	5	45.5	3261	4	09Y556	09y556 homo sapien	501	4	36.4	60	5	09VZL9	09vzl9 drosophila
429	5	45.5	3379	5	09V5N8	09v5n8 drosophila	502	4	36.4	60	5	09U3W3	09u3w3 caenorhabdi
430	5	45.5	3726	5	09VFL1	09vfl1 drosophila	503	4	36.4	61	2	09UTZ8	09utz8 nelisseria m
431	5	45.5	6875	6	028733	028733 oryctolagus	504	4	36.4	61	5	018358	018358 caenorhabdi
432	5	45.5	26926	4	010466	010466 homo sapien	505	4	36.4	62	1	09HL02	09hl02 thermoplasma
433	4	36.4	17	13	P81904	P81904 ictalurus p	506	4	36.4	62	2	047216	047216 escherichia
434	4	36.4	19	2	09R524	09r524 clostridium	507	4	36.4	62	5	017536	017536 caenorhabdi
435	4	36.4	20	2	09R424	09r424 proteus mir	508	4	36.4	62	14	010289	010289 oryza pseu
436	4	36.4	20	13	P81903	P81903 ictalurus p	509	4	36.4	63	5	016829	016829 drosophila
437	4	36.4	21	5	09TWF7	09twf7 entamoeba h	510	4	36.4	63	13	090951	090951 gallus gall
438	4	36.4	22	5	09TW08	09tw08 crithidia f	511	4	36.4	63	14	085040	085040 pseudodrabie
439	4	36.4	22	13	09PS65	09ps65 oncorhynch	512	4	36.4	64	2	045170	045170 borrelia ga
440	4	36.4	24	10	09S825	09s825 spinacia ol	513	4	36.4	64	2	09JXF6	09jxf6 nelisseria m
441	4	36.4	25	2	09R5H6	09r5h6 bacillus su	514	4	36.4	64	2	09JWJ3	09jwj3 nelisseria m
442	4	36.4	32	14	084265	084265 human papil	515	4	36.4	64	2	091026	091026 pseudomonas
443	4	36.4	34	2	09R5B3	09r5b3 streptomyce	516	4	36.4	64	10	09LGP2	09lgp2 oryza sativ
444	4	36.4	36	2	09K574	09k574 vibrio chol	517	4	36.4	65	5	09W5R0	09w5r0 drosophila
445	4	36.4	37	10	092P81	092p81 raphanus sa	518	4	36.4	66	2	09JTC8	09jtc8 nelisseria m
446	4	36.4	38	5	024462	024462 nitox novae	519	4	36.4	66	14	084302	084302 human papil
447	4	36.4	38	8	079167	079167 nitox novae	520	4	36.4	67	2	P76099	P76099 escherichia
448	4	36.4	38	8	079172	079172 nitox novae	521	4	36.4	67	3	014391	014391 schizosacch
449	4	36.4	39	4	09U0N1	09u0n1 homo sapien	522	4	36.4	67	4	012975	012975 homo sapien
450	4	36.4	39	4	09U0N1	09u0n1 homo sapien	523	4	36.4	68	2	049119	049119 methylodact
451	4	36.4	41	2	031327	031327 bacillus br	524	4	36.4	69	1	09HQE8	09hqe8 halobacteri
452	4	36.4	42	2	09X8F1	09x8f1 streptomyce	525	4	36.4	70	2	056716	056716 vibrio para
453	4	36.4	42	11	09ZIC3	09zic3 mus pretus	526	4	36.4	70	4	09NTH2	09nth2 homo sapien
454	4	36.4	43	4	013063	013063 homo sapien	527	4	36.4	70	5	09N7C7	09n7c7 leishmania
455	4	36.4	43	4	014909	014909 homo sapien	528	4	36.4	70	14	065909	065909 human coxa
456	4	36.4	43	5	09VD93	09vd93 drosophila	529	4	36.4	71	2	09ZFU3	09zfu3 mastigoclad
457	4	36.4	45	14	09Q1F1	09q1f1 polyomaviru	530	4	36.4	71	2	09RSM1	09rsm1 delnecoccus

239	5	45.5	528	10	022549	022549	Lycopersico	312	5	45.5	859	4	09GR7	09gr7 homo sapien
240	5	45.5	534	5	044117	044117	dtrosophila	313	5	45.5	866	2	09KR26	09kr26 streptococ
241	5	45.5	536	1	09HPA2	09HPA2	halobacteri	314	5	45.5	869	4	09HCF3	09hcf3 homo sapien
242	5	45.5	537	2	046977	046977	eschericlia	315	5	45.5	876	2	093576	093576 synechococ
243	5	45.5	540	10	048683	048683	arabidopsis	316	5	45.5	883	2	09X5P4	09x5p4 pseudomonas
244	5	45.5	545	11	060626	060626	mus musculu	317	5	45.5	891	2	09HYJ7	09hyj7 pseudomonas
245	5	45.5	550	5	09VN76	09VN76	dtrosophila	318	5	45.5	891	5	09V172	09v172 dtrosophila
246	5	45.5	559	5	016124	016124	dtrosophila	319	5	45.5	905	3	09P8C5	09p8c5 tolypoccladi
247	5	45.5	569	10	P93543	P93543	sambucus n1	320	5	45.5	922	11	09QXJ2	09qxj2 mus musculu
248	5	45.5	576	2	031745	031745	bacillus su	321	5	45.5	923	11	09WVL2	09wvl2 mus musculu
249	5	45.5	576	4	09NTE7	09NTE7	homo sapien	322	5	45.5	925	11	09QZB4	09qzb4 mus musculu
250	5	45.5	578	3	09V7N0	09V7N0	schizosacch	323	5	45.5	927	5	021811	021811 caenorhabdi
251	5	45.5	579	4	09PI73	09PI73	homo sapien	324	5	45.5	932	5	09NH67	09nh67 dtrosophila
252	5	45.5	581	2	09RU45	09RU45	delinococcus	325	5	45.5	932	5	09VAM0	09vam0 dtrosophila
253	5	45.5	581	10	09RTV1	09RTV1	canavalia 1	326	5	45.5	934	4	09P2B0	09p2b0 homo sapien
254	5	45.5	583	2	09Z5K5	09Z5K5	mycobacteri	327	5	45.5	943	2	04I09	04i09 actinomyc
255	5	45.5	583	5	019275	019275	caenorhabdi	328	5	45.5	950	5	09VF01	09vf01 dtrosophila
256	5	45.5	587	2	044563	044563	anabaena va	329	5	45.5	968	11	086522	086522 mus musculu
257	5	45.5	590	5	09GUG6	09GUG6	caenorhabdi	330	5	45.5	971	5	09TVL9	09tvl9 caenorhabdi
258	5	45.5	593	2	051191	051191	neisseria m	331	5	45.5	972	11	09OXM2	09oxm2 ratius norv
259	5	45.5	601	5	09W3V9	09W3V9	dtrosophila	332	5	45.5	973	4	09P203	09p203 homo sapien
260	5	45.5	610	14	084595	084595	paramecium	333	5	45.5	986	2	09PJT3	09pj3 chlamydia m
261	5	45.5	611	4	09NOE0	09NOE0	homo sapien	334	5	45.5	993	4	09HCC9	09hcc9 homo sapien
262	5	45.5	613	3	005785	005785	saccharomyc	335	5	45.5	993	5	017162	017162 brugia mala
263	5	45.5	613	4	09UFB1	09UFB1	homo sapien	336	5	45.5	1012	11	064028	064028 mus musculu
264	5	45.5	614	5	P91720	P91720	dtrosophila	337	5	45.5	1013	11	09JTH3	09jth3 ratius norv
265	5	45.5	622	2	031248	031248	actinobact	338	5	45.5	1028	5	09N5Y8	09n5y8 caenorhabdi
266	5	45.5	623	13	057602	057602	ictalurus p	339	5	45.5	1028	10	09RFH7	09rfh7 arabidopsis
267	5	45.5	626	5	09W1U4	09W1U4	dtrosophila	340	5	45.5	1039	5	09VTR6	09vtr6 dtrosophila
268	5	45.5	632	2	09KM53	09KM53	pectobacter	341	5	45.5	1040	4	09NTJ4	09ntj4 homo sapien
269	5	45.5	648	10	09M8Z4	09M8Z4	arabidopsis	342	5	45.5	1041	2	09JMD5	09jmd5 neisseria m
270	5	45.5	650	2	031382	031382	bradyrhizob	343	5	45.5	1049	3	013344	013344 emericella
271	5	45.5	652	4	09RK05	09RK05	streptomyce	344	5	45.5	1052	4	060264	060264 homo sapien
272	5	45.5	652	4	015410	015410	homo sapien	345	5	45.5	1062	4	09UL64	09ul64 homo sapien
273	5	45.5	654	2	039656	039656	plectonema	346	5	45.5	1072	5	09N5Y6	09n5y6 caenorhabdi
274	5	45.5	659	10	09LMU4	09LMU4	arabidopsis	347	5	45.5	1077	5	09J217	09j217 leishmania
275	5	45.5	667	2	09WXA5	09WXA5	erwinia car	348	5	45.5	1098	5	050733	050733 borrelia bu
276	5	45.5	668	10	09FUZ7	09FUZ7	zea mays (m	349	5	45.5	1105	4	09ULP3	09ulp3 homo sapien
277	5	45.5	680	2	052644	052644	ruminococcu	350	5	45.5	1127	2	09P843	09p843 mycobacteri
278	5	45.5	695	11	09ESB0	09ESB0	rattus norv	351	5	45.5	1132	5	002432	002432 dtrosophila
279	5	45.5	697	10	09FTG1	09FTG1	arabidopsis	352	5	45.5	1151	5	044319	044319 anurida mar
280	5	45.5	711	4	09ULG6	09ULG6	homo sapien	353	5	45.5	1151	13	057580	057580 gallus gall
281	5	45.5	718	2	09JZD0	09JZD0	neisseria m	354	5	45.5	1177	5	021391	021391 caenorhabdi
282	5	45.5	722	5	09NKT5	09NKT5	leishmania	355	5	45.5	1178	5	09VCC5	09vc5 dtrosophila
283	5	45.5	722	13	09PSZ5	09PSZ5	lampetra ja	356	5	45.5	1182	5	024352	024352 dtrosophila
284	5	45.5	728	5	09XGX4	09XGX4	caenorhabdi	357	5	45.5	1198	4	09H9J4	09h9j4 homo sapien
285	5	45.5	730	2	09X6U0	09X6U0	streptococ	358	5	45.5	1215	5	09V787	09v787 dtrosophila
286	5	45.5	733	10	09ZVD0	09ZVD0	arabidopsis	359	5	45.5	1216	11	09Z0Y7	09z0y7 mus musculu
287	5	45.5	749	4	09UK04	09UK04	homo sapien	360	5	45.5	1224	13	09Y190	09y190 xenopus lae
288	5	45.5	756	4	09Y6W2	09Y6W2	homo sapien	361	5	45.5	1228	2	044334	044334 agrobacteri
289	5	45.5	756	4	09Y6W1	09Y6W1	homo sapien	362	5	45.5	1236	11	070349	070349 mus musculu
290	5	45.5	756	4	09NRK8	09NRK8	homo sapien	363	5	45.5	1237	13	091976	091976 gallus gall
291	5	45.5	768	4	09HVA6	09HVA6	pseudomonas	364	5	45.5	1238	5	09V8B6	09v8b6 dtrosophila
292	5	45.5	768	11	088797	088797	rattus norv	365	5	45.5	1249	5	09N8G4	09n8g4 trypanosoma
293	5	45.5	770	4	013598	013598	homo sapien	366	5	45.5	1257	2	P96746	P96746 corneabacte
294	5	45.5	774	5	094892	094892	dtrosophila	367	5	45.5	1262	5	09W110	09w110 dtrosophila
295	5	45.5	779	4	09Y2F4	09Y2F4	homo sapien	368	5	45.5	1266	2	09W4Q4	09w4q4 dtrosophila
296	5	45.5	782	4	09NKR3	09NKR3	homo sapien	369	5	45.5	1295	5	09F6T1	09f6t1 escherichia
297	5	45.5	782	10	041583	041583	triticum ae	370	5	45.5	1329	10	09ZUK1	09zuk1 arabidopsis
298	5	45.5	792	3	09L2I8	09L2I8	streptomyce	371	5	45.5	1331	4	09UHC6	09uhc6 homo sapien
299	5	45.5	797	3	09P7M6	09P7M6	schizosacch	372	5	45.5	1339	4	09U012	09u012 homo sapien
300	5	45.5	800	2	09EXG2	09EXG2	listeria mo	373	5	45.5	1356	5	021946	021946 caenorhabdi
301	5	45.5	806	10	09MBT5	09MBT5	arabidopsis	374	5	45.5	1376	5	09W5D0	09w5d0 dtrosophila
302	5	45.5	806	10	09LPE0	09LPE0	arabidopsis	375	5	45.5	1400	11	09ESU6	09esu6 mus musculu
303	5	45.5	820	5	021737	021737	caenorhabdi	376	5	45.5	1403	5	09U6A2	09u6a2 dtrosophila
304	5	45.5	832	2	09KXZ7	09KXZ7	streptomyce	377	5	45.5	1409	5	002626	002626 caenorhabdi
305	5	45.5	832	5	09USN8	09USN8	leishmania	378	5	45.5	1416	5	019378	019378 caenorhabdi
306	5	45.5	833	5	09K3E2	09K3E2	streptomyce	379	5	45.5	1435	4	09UEB9	09ueb9 homo sapien
307	5	45.5	835	2	09NTN9	09NTN9	homo sapien	380	5	45.5	1435	4	09HCD7	09hcd7 homo sapien
308	5	45.5	838	4	09Y8B5	09Y8B5	cryptococcu	381	5	45.5	1445	4	09ULI6	09uli6 homo sapien
309	5	45.5	841	3	09PAQ2	09PAQ2	pneumocysti	382	5	45.5	1520	2	09KTO4	09ktq4 vibrio chol
310	5	45.5	852	3				383	5	45.5	1543	5	09VMA0	09vma0 dtrosophila
311	5	45.5						384	5	45.5				

93	5	45.5	212	14	09WH18	09wh18 fish lympho	166	5	45.5	373	10	P93031	P93031 arbidopsi
94	5	45.5	217	5	09VL41	09vl41 drosophila	167	5	45.5	373	10	09SD30	09sd30 arbidopsi
95	5	45.5	218	5	09PF65	09pf65 xylella fas	168	5	45.5	375	10	09RW21	09rw21 deinococc
96	5	45.5	220	5	09VBE1	09vbe1 drosophila	169	5	45.5	375	4	014712	014712 homo sapien
97	5	45.5	223	5	021648	021648 caenorhabd	170	5	45.5	376	13	09FWM1	09fwm1 gallus gall
98	5	45.5	224	3	042710	042710 schizosacch	171	5	45.5	381	5	09NDP7	09ndp7 clona lites
99	5	45.5	224	13	093590	093590 xenopus lae	172	5	45.5	382	10	003990	003990 daucus caro
100	5	45.5	225	5	001833	001833 caenorhabd	173	5	45.5	383	2	09X6L5	09x6l5 klebsiella
101	5	45.5	228	10	09ZPK9	09zpk9 hyacinthus	174	5	45.5	383	2	001678	001678 pneumocyst
102	5	45.5	229	3	09USP6	09usp6 schizosacch	175	5	45.5	384	2	056617	056617 vibrio chol
103	5	45.5	241	5	023790	023790 chironomus	176	5	45.5	385	5	020103	020103 caenorhabd
104	5	45.5	241	5	017321	017321 chironomus	177	5	45.5	387	14	066407	066407 dashneen mos
105	5	45.5	241	5	09G0H1	09guh1 caenorhabd	178	5	45.5	388	5	023232	023232 caenorhabd
106	5	45.5	242	2	09EUQ7	09eug7 streptococ	179	5	45.5	395	14	069369	069369 cercopithe
107	5	45.5	243	5	023784	023784 chironomus	180	5	45.5	399	4	09H4K6	09h4k6 homo sapien
108	5	45.5	246	2	09EX05	09ex05 streptomyc	181	5	45.5	403	10	004772	004772 chlorococc
109	5	45.5	256	11	09ESA6	09esa6 rattus norv	182	5	45.5	404	4	075177	075177 homo sapien
110	5	45.5	258	4	09NUD0	09nud0 homo sapien	183	5	45.5	405	2	09FCM7	09fcm7 streptomyc
111	5	45.5	258	13	091928	091928 bambusicola	184	5	45.5	405	13	091929	091929 bambusicola
112	5	45.5	262	5	015758	015758 trypanosoma	185	5	45.5	408	2	067062	067062 aquilex aeo
113	5	45.5	262	5	09V335	09v335 drosophila	186	5	45.5	412	1	09YBC6	09ybc6 aeropyrum p
114	5	45.5	262	5	09U396	09u396 caenorhabd	187	5	45.5	412	5	008523	008523 ascaris suu
115	5	45.5	263	5	09V334	09v334 drosophila	188	5	45.5	415	10	09FYP4	09fyp4 oryza saliv
116	5	45.5	267	5	09V6B8	09v6b8 drosophila	189	5	45.5	416	9	P90878	P90878 caenorhabd
117	5	45.5	269	2	09KPI3	09kpi3 vibrio chol	190	5	45.5	416	9	09ZMW9	09zwm9 corynephage
118	5	45.5	269	2	09F2T4	09f2t4 streptomyc	191	5	45.5	423	5	09VXB3	09vxb3 drosophila
119	5	45.5	270	2	P73854	P73854 synechocyst	192	5	45.5	424	3	074211	074211 yarrowia ll
120	5	45.5	270	2	09KAR9	09kar9 bacillus ha	193	5	45.5	425	1	013358	013358 homo sapien
121	5	45.5	270	2	09FAC6	09fac6 pseudomonas	194	5	45.5	431	4	09U731	09u731 pyrococcus
122	5	45.5	272	10	09FFV7	09ffv7 arbidopsi	195	5	45.5	434	3	09P970	09p970 tricholoma
123	5	45.5	273	2	09L1R2	09l1r2 streptomyc	196	5	45.5	437	3	09HEX2	09hex2 pneumocyst
124	5	45.5	274	13	091927	091927 bambusicola	197	5	45.5	437	5	09GVE6	09gve6 leishmania
125	5	45.5	274	13	091926	091926 bambusicola	198	5	45.5	439	5	09TX40	09tx40 dictyostell
126	5	45.5	276	10	09SNS1	09sns1 oryza sativ	199	5	45.5	442	5	026909	026909 trypanosoma
127	5	45.5	279	5	015759	015759 trypanosoma	200	5	45.5	444	5	026895	026895 trypanosoma
128	5	45.5	280	2	P95727	P95727 streptomyc	201	5	45.5	445	2	066913	066913 aquilex aeo
129	5	45.5	289	14	098505	098505 paramedium	202	5	45.5	445	2	09KBS1	09kbs1 bacillus ha
130	5	45.5	296	5	018675	018675 caenorhabd	203	5	45.5	448	2	09REQ3	09req3 zymomonas m
131	5	45.5	297	5	09V5R4	09v5r4 drosophila	204	5	45.5	448	5	023571	023571 caenorhabd
132	5	45.5	298	11	09ESA9	09esa9 rattus norv	205	5	45.5	448	10	023100	023100 arbidopsi
133	5	45.5	301	10	09ZPH0	09zph0 arbidopsi	206	5	45.5	451	2	09K111	09k111 neisseria m
134	5	45.5	302	5	09U719	09u719 chaetopteru	207	5	45.5	451	2	09JSW6	09jsw6 neisseria m
135	5	45.5	303	2	09RS08	09rs08 deinococc	208	5	45.5	455	10	09LJH2	09ljh2 arbidopsi
136	5	45.5	304	2	09URJ3	09urj3 neisseria m	209	5	45.5	456	2	09ZC76	09zc76 yersinia pe
137	5	45.5	306	5	09VRA8	09vra8 drosophila	210	5	45.5	458	5	09N3U1	09n3u1 caenorhabd
138	5	45.5	307	10	09ZUP7	09zup7 arbidopsi	211	5	45.5	461	2	09LAX6	09lax6 streptococ
139	5	45.5	315	5	P91282	P91282 caenorhabd	212	5	45.5	471	3	09P723	09p723 neutrospora
140	5	45.5	320	5	09VQ29	09vq29 drosophila	213	5	45.5	476	5	003996	003996 plasmodium
141	5	45.5	321	2	084847	084847 chlamydia t	214	5	45.5	479	2	09H2B7	09h2b7 pseudomonas
142	5	45.5	321	2	09R705	09r705 escherichia	215	5	45.5	480	5	09VMD5	09vmd5 drosophila
143	5	45.5	323	2	045621	045621 bacillus sp	216	5	45.5	480	10	004535	004535 arbidopsi
144	5	45.5	326	3	09P958	09p958 tricholoma	217	5	45.5	485	2	09ZNM1	09znm1 pseudomonas
145	5	45.5	326	3	09P957	09p957 tricholoma	218	5	45.5	489	10	065518	065518 arbidopsi
146	5	45.5	327	2	09Z9U8	09z9u8 bacillus ha	219	5	45.5	489	11	09W0H5	09wh05 mus musculi
147	5	45.5	330	5	018118	018118 caenorhabd	220	5	45.5	500	5	P90972	P90972 caenorhabd
148	5	45.5	330	14	09YTN3	09ytn3 ateline her	221	5	45.5	500	5	093289	093289 caenorhabd
149	5	45.5	338	2	09HVA2	09hva2 pseudomonas	222	5	45.5	501	11	070399	070399 cavia porce
150	5	45.5	340	2	09PHN5	09phn5 campylobact	223	5	45.5	502	5	09V823	09v823 drosophila
151	5	45.5	345	4	09NRD1	09nrd1 homo sapien	224	5	45.5	503	5	09VLA8	09vla8 drosophila
152	5	45.5	347	2	005160	005160 synechococ	225	5	45.5	506	4	09NMW9	09nmw9 homo sapien
153	5	45.5	347	13	09DFG2	09dfg2 xenopus lae	226	5	45.5	509	5	09GUK1	09guk1 leishmania
154	5	45.5	349	4	015413	015413 homo sapien	227	5	45.5	510	5	015767	015767 babesia big
155	5	45.5	350	5	015760	015760 babesia big	228	5	45.5	511	5	09VVS5	09vvs5 drosophila
156	5	45.5	350	5	015761	015761 babesia big	229	5	45.5	511	10	09MB73	09mb73 citrus unsh
157	5	45.5	352	2	09XAI4	09xai4 streptomyc	230	5	45.5	513	2	054044	054044 pseudomonas
158	5	45.5	354	5	026635	026635 sarcophaga	231	5	45.5	513	10	09FVBB	09fvbb brassica na
159	5	45.5	357	2	053459	053459 mycobacteri	232	5	45.5	515	5	09TX44	09tx44 drosophila
160	5	45.5	364	5	002429	002429 dysdera cro	233	5	45.5	520	4	09H239	09h239 homo sapien
161	5	45.5	369	2	09RPE4	09rpe4 mycoplasma	234	5	45.5	520	10	09FVBC	09fvbc brassica na
162	5	45.5	371	5	046169	046169 tribolium c	235	5	45.5	520	10	09FVBB	09fvbb brassica na
163	5	45.5	371	5	046052	046052 drosophila	236	5	45.5	524	5	09VFB9	09vfb9 brassica na
164	5	45.5	372	5	09VXY1	09vxy1 drosophila	237	5	45.5	527	5	09VAB7	09vab7 drosophila
165	5	45.5	373	2	09PB07	09pb07 xylella fas	238	5	45.5	527	10	041479	041479 solanum tub

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OK protein - protein search, using sw model

Run on: August 15, 2001, 12:42:32 ; Search time 58.01 Seconds  
(without alignments)  
25.088 Million cell updates/sec

Title: US-09-372-036-29  
Perfect score: 11  
Sequence: 1 OQQTAPKAPTE 11

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL\_16:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_protist:\*

12: sp\_unclassified:\*

13: sp\_vertebrate:\*

14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110.0	478	2	003493	Q03493 listeria mo
2	100.9	76	2	048784	Q48784 listeria mo
3	100.9	84	2	057205	Q57205 listeria mo
4	100.9	86	2	048775	Q48775 listeria mo
5	100.9	86	2	048782	Q48782 listeria mo
6	100.9	88	2	048781	Q48781 listeria mo
7	100.9	117	2	048774	Q48774 listeria mo
8	90.9	236	5	09NAN3	Q9NAN3 caenorhabdi
9	54.5	288	5	09EMH8	Q9EMH8 streptomyc
10	54.5	329	2	09RGS0	Q9RGS0 staphylococ
11	54.5	343	5	019781	Q19781 caenorhabdi
12	54.5	348	10	09FPA4	Q9FPA4 oryza sativ
13	54.5	371	5	020966	Q20966 caenorhabdi
14	54.5	388	5	09VMV9	Q9VMV9 drosophila
15	54.5	606	2	086584	Q86584 streptomyc
16	54.5	612	2	092614	Q92614 streptomyc
17	54.5	663	2	09P196	Q9P196 campylobact
18	54.5	672	5	025586	Q25586 onchocerca
19	54.5	689	2	09RKV7	Q9RKV7 streptomyc

20	54.5	855	5	09NEW7	Q9NEW7 caenorhabdi
21	54.5	861	2	09HUA6	Q9HUA6 pseudomonas
22	54.5	925	5	09YEW6	Q9YEW6 drosophila
23	54.5	995	5	P90897	P90897 caenorhabdi
24	54.5	1052	11	088332	Q88332 mus musculu
25	54.5	1158	5	097169	Q97169 drosophila
26	54.5	1199	5	09V5G1	Q9V5G1 drosophila
27	54.5	1241	3	013617	Q13617 schistosac
28	54.5	1257	5	09NMF04	Q9NMF04 caenorhabdi
29	54.5	2121	5	09NH54	Q9NH54 drosophila
30	54.5	2129	5	09VLT23	Q9VLT23 drosophila
31	54.5	2478	5	09VXZ5	Q9VXZ5 drosophila
32	54.5	5233	5	09NB71	Q9NB71 drosophila
33	45.5	22	2	P81150	P81150 desulfovibr
34	45.5	23	2	P81149	P81149 desulfovibr
35	45.5	61	2	056859	Q56859 yersinia en
36	45.5	64	2	Q9S1E1	Q9S1E1 yersinia pe
37	45.5	64	2	Q9JP69	Q9JP69 escherichia
38	45.5	82	10	Q9MAC9	Q9MAC9 hordium vul
39	45.5	88	13	Q9YCP4	Q9YCP4 gallus gall
40	45.5	95	10	Q38792	Q38792 antirrhinum
41	45.5	103	2	Q91063	Q91063 streptomyc
42	45.5	110	5	Q9N3V6	Q9N3V6 caenorhabdi
43	45.5	111	11	Q9ESA8	Q9ESA8 ratius norv
44	45.5	114	4	Q95014	Q95014 homo sapien
45	45.5	114	5	Q9VBH2	Q9VBH2 drosophila
46	45.5	117	5	Q9ULN7	Q9ULN7 leishmania
47	45.5	118	2	Q9RTX9	Q9RTX9 delinococcus
48	45.5	120	3	Q9PBW7	Q9PBW7 pneumocysti
49	45.5	121	2	Q9PP33	Q9PP33 campylobact
50	45.5	124	13	Q9PS90	Q9PS90 gallus gall
51	45.5	127	1	Q9Y8T9	Q9Y8T9 aeropyrum p
52	45.5	133	14	Q9JSE3	Q9JSE3 fowlpox vir
53	45.5	136	2	Q9K1X8	Q9K1X8 vibrio chol
54	45.5	136	11	Q9ESA7	Q9ESA7 ratius norv
55	45.5	137	5	Q03997	Q03997 plasmodium
56	45.5	137	10	Q40043	Q40043 hordium vul
57	45.5	139	9	P93179	P93179 hordium vul
58	45.5	144	2	Q9KZNI	Q9KZNI streptomyc
59	45.5	144	2	Q9JPP4	Q9JPP4 neisseria m
60	45.5	146	3	Q9HG53	Q9HG53 pneumocysti
61	45.5	146	3	Q9HG52	Q9HG52 pneumocysti
62	45.5	146	3	Q9HG51	Q9HG51 pneumocysti
63	45.5	148	2	Q9K173	Q9K173 neisseria m
64	45.5	156	10	Q64940	Q64940 lophopyrum
65	45.5	157	5	Q9YNB9	Q9YNB9 drosophila
66	45.5	158	10	Q41600	Q41600 triticum tu
67	45.5	159	2	Q9JTC9	Q9JTC9 neisseria m
68	45.5	159	4	Q9H8U1	Q9H8U1 homo sapien
69	45.5	166	2	Q44075	Q44075 aeromonas h
70	45.5	170	4	Q9H4W2	Q9H4W2 homo sapien
71	45.5	172	4	Q9P054	Q9P054 homo sapien
72	45.5	172	5	Q9N985	Q9N985 leishmania
73	45.5	174	5	Q95992	Q95992 pyrococcus
74	45.5	175	11	Q63571	Q63571 ratius norv
75	45.5	176	3	Q9P765	Q9P765 neurospora
76	45.5	176	10	Q9ST24	Q9ST24 zeia mays (m
77	45.5	177	10	Q9ZWL5	Q9ZWL5 zeia mays (m
78	45.5	181	2	Q05816	Q05816 mycobacteri
79	45.5	181	4	Q9Y6Z4	Q9Y6Z4 homo sapien
80	45.5	181	5	Q9VBZ0	Q9VBZ0 drosophila
81	45.5	183	2	Q9RWB9	Q9RWB9 delinococcus
82	45.5	191	5	Q18649	Q18649 caenorhabdi
83	45.5	194	11	Q9JUK2	Q9JUK2 mus musculu
84	45.5	197	2	Q9RC54	Q9RC54 bacillus ha
85	45.5	197	5	Q9YV38	Q9YV38 drosophila
86	45.5	200	2	Q9F459	Q9F459 anabaena sp
87	45.5	202	3	Q9HGF8	Q9HGF8 saccharomyc
88	45.5	205	2	Q9JYB8	Q9JYB8 neisseria m
89	45.5	207	2	Q9KAD3	Q9KAD3 bacillus ha
90	45.5	207	10	Q9LY40	Q9LY40 arabidopsis
91	45.5	208	4	Q75852	Q75852 homo sapien
92	45.5	212	2	Q9RSN7	Q9RSN7 delinococcus

**THIS PAGE BLANK (uspto)**

---

```

RX MEDLINE-92347850; PubMed-1386335;
RA Inaba T., Matsushima H., Valentine M., Rousset M.F., Sherr C.J.,
RA Look A.T.;
RT "Genomic organization, chromosomal localization, and independent
RT expression of human cyclin D genes.";
RL Genomics 13:565-574(1992).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE Holoenzyme complex. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M90814; AAS1927.1; -
DR EMBL: M92287; AAS52337.1; -
DR EMBL: M88087; AAS1929.1; -
DR EMBL: M88084; AAS1929.1; JOINED.
DR EMBL: M88085; AAS1929.1; JOINED.
DR EMBL: M88086; AAS1929.1; JOINED.
DR PIR: BA2822; B42822.
DR PIR: A44022; A44022.
DR MIM: 123834; -
DR InterPro: IPR000553; -
DR Pfam: PF00134; cyclin: 1.
DR PROSITE: PS00292; CYCLINS: 1.
KW Cyclin; Cell cycle; Cell division; Multigene family.
FT CONFLICT 259 259 A -> S (IN REF. 2).
SQ SEQUENCE 292 AA; 32503 MW; 16FE56FE2FE8B029 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 292;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9
DB 266 APKAP 270

RESULT 50
SMDF_HUMAN STANDARD; PRT; 296 AA.
AC 015491;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE NEUREGULIN-1, SENSORY AND MOTOR NEURON-DERIVED FACTOR ISOFORM.
GN NRGI OR HGL OR NDF OR HRGA OR GGF OR SMDF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain stem, and Cerebellum;
RX MEDLINE-95301541; PubMed-7782315;
RA Ho W.-H., Armanini M.P., Nijens A., Phillips H.S., Osheroff P.L.;
RT "Sensory and motor neuron-derived factor. A novel heregulin variant
RT highly expressed in sensory and motor neurons.";
RL J. Biol. Chem. 270:14523-14532(1995).
CC -1- FUNCTION: THE ISOFORM SMDF MAY PLAY A ROLE IN MOTOR AND SENSORY
CC NEURON DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: SECRETED. MAY POSSESS AN INTERNAL UNCLEAVED
CC SIGNAL SEQUENCE.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 10 ISOFORMS OF NRGI ARE PRODUCED BY

```

```

CC ALTERNATIVE SPLICING. EXCEPT FOR SMDF THEY ARE IN ENTRY AC 002297.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL CORD MOTOR
CC NEURONS, DORSAL ROOT GANGLION NEURONS, AND BRAIN. PREDOMINANT
CC ISOFORM EXPRESSED IN SENSORY AND MOTOR NEURONS. NOT DETECTED IN
CC ADULT HEART, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, AND
CC PANCREAS. NOT EXPRESSED IN FETAL LUNG, LIVER, AND KIDNEY.
CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED IN DEVELOPING SPINAL MOTOR
CC NEURONS AND IN DEVELOPING CRANIAL NERVE NUCLEI. EXPRESSION IS
CC MAINTAINED ONLY IN BOTH ADULT MOTOR NEURONS AND DORSAL ROOT
CC GANGLION NEURONS.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: I41827; AAC1764.1; -
DR MIM: 142445; -
DR InterPro: IPR000561; -
DR Pfam: PF00008; EGF_1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; FALSE-NEG.
KW Growth factor; EGF-like domain; Transmembrane; Multigene family;
KW Alternative splicing.
FT TRANSMEM 76 100
FT DOMAIN 58 91 INTERNAL SIGNAL SEQUENCE (POTENTIAL).
FT DOMAIN 211 232 CYS-RICH.
FT DOMAIN 233 277 SER/THR-RICH.
FT DISULFD 237 251 EGF-LIKE.
FT DISULFD 245 265 BY SIMILARITY.
FT DISULFD 267 276 BY SIMILARITY.
SQ SEQUENCE 296 AA; 31685 MW; 8DA1743217F7EB02 CRC64;

```

```

Query Match 45.5%; Score 5; DB 1; Length 296;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 OTAPK 7
DB 217 OTAPK 221

```

Search completed: August 15, 2001, 12:43:01  
Job time: 491 sec



```

LBX1_HUMAN
ID LBX1_HUMAN STANDARD; PRT; 280 AA.
AC P52954;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE TRANSCRIPTION FACTOR LBX1.
GN LBX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=96227534; PubMed=8645601;
RA Jagla K., Delle P., Mattei M.-G., Jagla T., Schunbaur B.,
RA Dretzen G., Bellard F., Bellard M.,
RT "Mouse Lbx1 and human LBX1 define a novel mammalian homeobox gene
RL family related to the Drosophila lady bird genes."
RL Mech. Dev. 53:345-356(1995).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X90828; CA62342.1; -.
CC DR HSSP; P22808; 1VND.
CC DR InterPro; IPR001356; -.
CC DR Pfam; PF00046; homeobox; 1.
CC DR PRINTS; PR00024; HOMEBOX_1.
CC DR PROSITE; PS00027; HOMEBOX_1; 1.
CC DR PROSITE; PS50071; HOMEBOX_2; 1.
CC KW Homeobox; DNA-binding; Transcription regulation; Nuclear protein.
CC FT DNA_BIND 124 183 HOMEBOX.
CC FT DOMAIN 219 225 POLY-GLY.
CC FT DOMAIN 269 280 ASP/GLU-RICH (HIGHLY ACIDIC).
CC SQ SEQUENCE 280 AA; 30307 MW; 8DDA14BDD7A40B97 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 280;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9
Db 240 APRAP 244

RESULT 48
LIPH_PSEAE STANDARD; PRT; 288 AA.
ID LIPH_PSEAE
AC Q01725;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LIPH PROTEIN.
GN LIPH OR PA2863.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAOI;
RX MEDLINE=92381478; PubMed=1512563;
RA Wohlfarth S., Hoesche C., Strunk C., Winkler U.K.;
RT "Molecular genetics of the extracellular lipase of Pseudomonas
RT aeruginosa PAOI."

```

```

RL J. Gen. Microbiol. 138:1325-1335(1992).
RN [2]
RN REVISION TO 250-280.
RA Wohlfarth S.,
RA Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Ralzer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
CC -1- FUNCTION: HELPS TO PRODUCE AN EXTRACELLULARLY ACTIVE LIPASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X63391; CAA44998.1; ALT_INIT.
CC DR EMBL; AE004712; AAG06251.1; -.
CC DR PIR; S25769; S25769.
CC SQ SEQUENCE 288 AA; 32440 MW; 8BF64F42B79CF94 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 288;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OQORA 5
Db 189 OQORA 193

RESULT 49
CGD3_HUMAN STANDARD; PRT; 292 AA.
ID CGD3_HUMAN
AC P30281;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE G1/S-SPECIFIC CYCLIN D3.
GN CCND3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93015922; PubMed=1383201;
RA Motokura T., Kiyomatsi K., Kronenberg H.M., Arnold A.;
RT "Cloning and characterization of human cyclin D3, a cDNA closely
RT related in sequence to the PRAD1/cyclin D1 proto-oncogene."
RL J. Biol. Chem. 267:20412-20415(1992).
RN [3]
RP SEQUENCE OF 52-237 FROM N.A.
RC Tissue=Placenta;

```

FT REPEAT 229 233 5.  
 FT REPEAT 235 239 6.  
 FT REPEAT 245 249 7.  
 SQ SEQUENCE 260 AA; 27715 MW; A6CE29582A626D10 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 260;  
 Best Local Similarity 100.0%; Pred. No. 56;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 DB 37 APKAP 41

RESULT 46  
 THEFT\_THEVU STANDARD; PRT; 279 AA.  
 AC P04072;  
 DT 01-NOV-1986 (Rel. 03, Created)  
 DT 01-NOV-1986 (Rel. 03, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE THERMITASE (EC 3.4.21.66).  
 OS Thermocactinomycetes vulgaris.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Thermocactinomycetes.  
 NX NCBI\_TaxID=2026;  
 RN [1]  
 RP SEQUENCE.  
 RA Meloun B., Baudys M., Kostka V., Hausdorf G., Frommel C., Hohné W.E.;  
 RT "Complete primary structure of thermitase from Thermocactinomycetes  
 RT vulgaris and its structural features related to the subtilisin-type  
 RT proteases.";  
 RL FEBS Lett. 183:195-200(1985).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RA MEDLINE=89171261; PubMed=2647518;  
 RA Teplyakov A.V., Kuranova I.P., Harutyunyan E.H., Frommel C.,  
 RA Hohné W.E.;  
 RT "Crystal structure of thermitase from Thermocactinomycetes vulgaris at  
 RT 2.2-A resolution.";  
 RL FEBS Lett. 244:208-212(1989).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (1.98 ANGSTROMS).  
 RA MEDLINE=90096158; PubMed=2689655;  
 RA Gros P., Betzel C., Dauter Z., Wilson K.S., Hol W.G.J.;  
 RT "Molecular dynamics refinement of a thermitase-eglin-c complex at  
 RT 1.96-A resolution and comparison of two crystal forms that differ in  
 RT calcium content.";  
 RL J. Mol. Biol. 210:347-367(1989).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).  
 RA MEDLINE=90317828; PubMed=2196375;  
 RA Teplyakov A.V., Kuranova I.P., Harutyunyan E.H., Vainshteln B.K.,  
 RA Frommel C., Hohné W.E., Wilson K.S.;  
 RT "Crystal structure of thermitase at 1.4-A resolution.";  
 RL J. Mol. Biol. 214:261-279(1990).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RA MEDLINE=91131592; PubMed=1993669;  
 RA Gros P., Kalk K.H., Hol W.G.J.;  
 RT "Calcium binding to thermitase. Crystallographic studies of  
 RT thermitase at 0, 5, and 100 mM calcium.";  
 RL J. Biol. Chem. 266:2953-2961(1991).  
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS, INCLUDING COLLAGEN.  
 CC -1- ENZYME REGULATION: BINDS TWO CALCIUM IONS WHICH SEEM TO PLAY AN  
 CC IMPORTANT ROLE IN THE STABILIZATION OF THE TERTIARY STRUCTURE OF  
 CC THE ENZYME.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE  
 CC SUBTILINASE FAMILY.  
 CC PIR: A00973; SUMYTV.  
 DR PDB: 1TEC; 15-OCT-89.  
 DR PDB: 2TEC; 15-JAN-92.

DR PDB: 3TEC; 15-JAN-92.  
 DR PDB: 1THM; 31-JAN-94.  
 DR MEROPS: S08.007; -;  
 DR InterPro: IPR000209; -;  
 DR Pfam: PF00082; Peptidase\_S8; 1.  
 DR PRINTS: PR00723; SUBTILISIN.  
 DR PROSITE: PS00136; SUBTILASE\_ASP; 1.  
 DR PROSITE: PS00137; SUBTILASE\_HIS; 1.  
 DR PROSITE: PS00138; SUBTILASE\_SER; 1.  
 KW Hydrolyase; Serine protease; Calcium-binding; 3D-structure.  
 FT ACT\_SITE 38 38  
 FT ACT\_SITE 71 71  
 FT ACT\_SITE 225 225  
 FT TURN 7 7  
 FT HELIX 8 11  
 FT HELIX 14 17  
 FT TURN 18 19  
 FT HELIX 20 24  
 FT TURN 25 26  
 FT TURN 30 31  
 FT STRAND 33 38  
 FT TURN 43 44  
 FT TURN 46 51  
 FT STRAND 52 57  
 FT TURN 58 61  
 FT STRAND 62 62  
 FT STRAND 71 80  
 FT HELIX 94 95  
 FT TURN 97 102  
 FT STRAND 106 107  
 FT TURN 112 124  
 FT HELIX 125 126  
 FT TURN 129 132  
 FT STRAND 136 136  
 FT STRAND 141 152  
 FT TURN 153 154  
 FT STRAND 156 160  
 FT STRAND 171 171  
 FT STRAND 172 174  
 FT TURN 176 177  
 FT STRAND 178 184  
 FT TURN 186 187  
 FT STRAND 190 190  
 FT TURN 192 193  
 FT TURN 198 199  
 FT STRAND 202 205  
 FT STRAND 209 213  
 FT TURN 214 216  
 FT STRAND 217 221  
 FT HELIX 224 239  
 FT TURN 240 242  
 FT HELIX 245 254  
 FT TURN 255 255  
 FT STRAND 257 257  
 FT TURN 260 261  
 FT STRAND 262 262  
 FT TURN 263 265  
 FT STRAND 266 266  
 FT STRAND 269 270  
 FT HELIX 273 278  
 SQ SEQUENCE 279 AA; 28366 MW; 8065049BC8927AC0 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 60;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
 DB 92 TAPKA 96

RESULT 47

```

RL Nature 393:537-544(1998).
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE - NAD(+) + UBIQUINOL.
CC -1- COFACTOR: BINDS A 2FE-2S CLUSTER (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 24 KDA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z83867; CAB06290.1; -.
CC Truncat: RV3149; -.
CC InterPro: IPR002023; -.
CC DR Pfam: PF01257; COMPLEX1_24kD.1.
CC DR PROSITE: PS01099; COMPLEX1_24K; FALSE_NEG.
CC KM Oxidoreductase; NAD; Ubiquinone; Iron-sulfur.
CC FT METAL 114 114 IRON-SULFUR (2FE-2S) (POTENTIAL).
CC FT METAL 119 119 IRON-SULFUR (2FE-2S) (POTENTIAL).
CC FT METAL 155 155 IRON-SULFUR (2FE-2S) (POTENTIAL).
CC FT METAL 159 159 IRON-SULFUR (2FE-2S) (POTENTIAL).
CC SEQUENCE 252 AA; 27197 MW; 07B62E1A282175A4 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 252;
Best Local Similarity 100.0%; Pred. No.55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PKAPT 10
   |||||
Db 192 PKAPT 196

RESULT 44
HXC9_FUGRU STANDARD; PRT; 259 AA.
ID AC 042502;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HOMEOBOX PROTEIN HOX-C9 (HOXC-9).
GN HOXC9.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE:97285126; PubMed-9140399;
RA Aparicio S., Hawker K., Collage A., Mikawa Y., Zuo L., Venkatesh B.,
RA Chen E., Krumlauf R., Brenner S.;
RT "Organization of the Fugu rubripes Hox clusters: evidence for
RT continuing evolution of vertebrate Hox complexes.";
RL Nat. Genet. 16:79-83(1997).
CC -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE ABD-B FAMILY OF HOMEOBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: U92572; AAB68680.1; -.
DR InterPro: IPR000047; -.
DR InterPro: IPR001356; -.
DR Pfam: PF000046; homeobox; 1.
DR PRINTS: PR00024; HOMEOBOX.
DR PRINTS: PR00031; HTHREPRESSR.
DR PROSITE: PS00027; HOMEOBOX_1; 1.
DR PROSITE: PS00071; HOMEOBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DNA_BIND 191 250 HOMEOBOX.
FT SEQUENCE 259 AA; 29691 MW; 1BF5ACB7558F15D2 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No.56;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APRAP 9
   |||||
Db 57 APRAP 61

RESULT 45
H11_VOLCA STANDARD; PRT; 260 AA.
ID AC 008654;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HISTONE H1-I.
GN H1-I.
OS Volvox carterii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-F. MAGARIENSIS / HK10;
RA MEDLINE:93328125; PubMed-8335260;
RA Lindauer A., Mueller K., Schmitt R.;
RT "Two histone H1-encoding genes of the green alga Volvox carterii with
RT features intermediate between plant and animal genes.";
RL Gene 129:59-68(1993).
CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF
CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS RESTRICTED TO EMBRYOGENESIS.
CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L07946; AAA74723.1; -.
CC PIR: JN0747; JN0747.
CC DR HSSP: P08287; IGHC.
CC DR InterPro: IPR001386; -.
CC DR Pfam: PF00538; linker_histone; 1.
KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family;
KW Repeat.
FT INIT_MET 0 0 BY SIMILARITY.
FT DOMAIN 55 132 GLOBULAR (BY SIMILARITY).
FT DNA_BIND 138 141 POTENTIAL.
FT DOMAIN 135 249 7 X 5 AA REPEATS OF K-K-[AS]-T-P.
FT REPEAT 135 139 1.
FT REPEAT 187 191 2.
FT REPEAT 198 202 3 (APPROXIMATE).
FT REPEAT 208 212 4.

```

Best Local Similarity 100.0%; Pred. No. 52;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 OTAPK 7  
DB 29 OTAPK 33

RESULT 41

NOCM\_AGRU ID NOCM\_AGRU STANDARD: PRT: 241 AA.

AC P35113.

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE NOLALINE TRANSPORT SYSTEM PERMEASE PROTEIN NOCM.

GN NOCM.

OS Agrobacterium tumefaciens.

OG Plasmid pR1C58.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OX NCBI\_TaxID=362;

RN NCBI\_TaxID=362;

RP SEQUENCE FROM N.A.

RX MEDLINE-92121122; PubMed-1732218;

RT Zanker H., von Lintig J., Schroeder J.;

RT "Optine transport genes in the octopine (oc) and nopaline (noc) catabolic regions in T1 plasmids of Agrobacterium tumefaciens."

RL J. Bacteriol. 174:841-849(1992).

CC -1- FUNCTION: COMPONENT OF THE NOLALINE ACTIVE TRANSPORT SYSTEM

CC IS ALSO CAPABLE OF TRANSPORTING OCTOPINE PROVIDED THAT CATABOLIC

CC FUNCTIONS ARE INDUCED WITH NOLALINE.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE

CC (BY SIMILARITY).

CC -1- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-

CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE HISMO

CC SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M77785; AAA50513.1; -.

DR PIR: F42600; F42600.

DR InterPro: IPR000515; -.

DR Pfam: PF00528; BPD\_transp.1.

DR PROSITE: PS00402; BPD\_TRANS\_PNN\_MEMB.1.

DR TRANSPORT: Plasmid; Transmembrane; Inner membrane.

FT TRANSMEM 13 41 POTENTIAL.

FT TRANSMEM 161 183 POTENTIAL.

FT TRANSMEM 194 218 POTENTIAL.

SEQUENCE 241 AA; 26424 MW; 390E7B6C070C5D65 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 241;

Best Local Similarity 100.0%; Pred. No. 52;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8

DB 233 TAPKA 237

RESULT 42

HIO\_CHITH

ID HIO\_CHITH STANDARD: PRT: 244 AA.

AC 007134;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE HISTONE H1, ORPHON.

OS Chironomus thummi thummi (Midge).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;

OC Chironomidae; Chironomidae; Chironominae; Chironomus.

OX NCBI\_TaxID=7155;

RN NCBI\_TaxID=7155;

RP SEQUENCE FROM N.A.

RX MEDLINE-94087747; PubMed-8263935;

RT Hankeln T., Schmidt E.R.;

RT "Divergent evolution of an 'orphan' histone gene cluster in

RT Chironomus."

RL J. Mol. Biol. 234:1301-1307(1993).

CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF

CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.

CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: X72803; CAAS1322.1; -.

DR HSP: P08287; IGHC.

DR InterPro: IPR001386; -.

DR Pfam: PF00538; Linker histone.1.

KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.

SEQUENCE 244 AA; 25455 MW; 12A6D32C6690354 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 244;

Best Local Similarity 100.0%; Pred. No. 53;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 ACPAP 9

DB 28 ACPAP 32

RESULT 43

NOOE\_MYCTU

ID NOOE\_MYCTU STANDARD: PRT: 252 AA.

AC P95177;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE NADH DEHYDROGENASE I CHAIN E (EC 1.6.5.3) (NADH-UBIQUINONE

DE OXIDOREDUCTASE CHAIN E).

GN NOOE OR RV3149 OR MYCY03A2.09C.

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1773;

RN NCBI\_TaxID=1773;

RP SEQUENCE FROM N.A.

RC STRAIN-H37RV;

RX MEDLINE-98295987; PubMed-9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Pettell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,

RA Taylor K., Whitehead S., Bartell B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the

complete genome sequence."

RX MEDLINE-92020152; PubMed-1923777;  
 RA Yang P., Katsura M., Nakayama T., Mikami K., Iwabuchi M.;  
 RT "Molecular cloning and nucleotide sequences of cDNAs for histone H1  
 and H2B variants from wheat.";  
 CC Nucleic Acids Res. 19:5077-5077(1991).  
 CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF  
 CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X59872; CAA42529.1; -  
 CC PIR; S22322; S22322.  
 CC HSSP; P02259; 1HST.  
 CC InterPro: IPR001386; -  
 CC Pfam: PF00538; linker\_histone; 1.  
 CC DR Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.  
 CC SQ SEQUENCE 236 AA; 24018 MW; 069105D1972C74BB CRC64;

Query Match 45.5%; Score 5; DB 1; Length 236;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 APRAP 9  
 Db 32 APRAP 36

RESULT 39  
 H1E.CHITE STANDARD; PRT; 237 AA.  
 AC P40278;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HISTONE H1E.  
 OS Chironomus tentans (Midge).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;  
 OC Chironomidae; Chironomidae; Chironominae; Chironomus.  
 CC NCBI\_TaxID=7153;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC RA Schulze E., Wisniewski J.R., Nagel S., Gavenis K., Grossbach U.;  
 CC Submitted (XX-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF  
 CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L29105; AAB53945.1; -  
 CC HSSP; P08287; IGHC.  
 CC InterPro: IPR001386; -  
 CC Pfam: PF00538; linker\_histone; 1.  
 CC DR Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.  
 CC SQ SEQUENCE 237 AA; 24689 MW; D429364FFBCB1F3E CRC64;

Query Match 45.5%; Score 5; DB 1; Length 237;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 APRAP 9  
 Db 28 APRAP 32

RESULT 40  
 OPAA.NEIGO STANDARD; PRT; 239 AA.  
 ID OPAA.NEIGO  
 AC 004876;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE OPACITY PROTEIN OPA53 PRECURSOR (FRAGMENT).  
 CC OPAA.  
 CC GN Neisseria gonorrhoeae.  
 CC OS Neisseria gonorrhoeae.  
 CC OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 CC NCBI\_TaxID=485;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC RP STRAIN-MS11 / F3;  
 CC RX MEDLINE-93178439; PubMed-8440254;  
 RA Kupsch E.-M., Knepper B., Kuroki T., Heuer I., Meyer T.F.;  
 RT "Variable opacity (Opa) outer membrane proteins account for the cell  
 RT tropisms displayed by Neisseria gonorrhoeae for human leukocytes and  
 RT epithelial cells.";  
 RL EMBO J. 12:641-650(1993).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 CC RP STRAIN-MS11 / V18;  
 CC RX MEDLINE-92114767; PubMed-1815562;  
 RA Bhat K.S., Gibbs C.P., Barrera O., Morrison S.G., Jaehnig F.,  
 RA Stern A., Kupsch E.-M., Meyer T.F., Swanson J.;  
 RT "The opacity proteins of Neisseria gonorrhoeae strain MS11 are  
 RT encoded by a family of 11 complete genes.";  
 RL Mol. Microbiol. 5:1889-1901(1991).  
 CC [3]  
 CC SEQUENCE FROM N.A.  
 CC RP ERRATUM.  
 CC RX MEDLINE-92261323; PubMed-1584024;  
 RA Bhat K.S., Gibbs C.P., Barrera O., Morrison S.G., Jaehnig F.,  
 RA Stern A., Kupsch E.-M., Meyer T.F., Swanson J.;  
 RL Mol. Microbiol. 6:1073-1076(1992).  
 CC -1- FUNCTION: IMPLICATED IN A NUMBER OF ADHERENCE FUNCTIONS. OPA  
 CC PROTEINS ARE IMPLICATED IN PATHOGENESIS AND ARE SUBJECT TO PHASE  
 CC VARIATION.  
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z18930; CAA79363.1; -  
 CC DR EMBL; X60709; CAA43119.1; -  
 CC PIR; S28630; S28630.  
 CC DR PIR; S36331; S36331.  
 CC KW Outer membrane; Multigene family; Signal.  
 CC FT NON\_TER 1  
 CC FT SIGNS 1  
 CC FT CHAIN 2  
 CC FT VARIANT 92 92  
 CC FT NON\_TER 239 239  
 CC SQ SEQUENCE 239 AA; 26734 MW; 2F12E69DEB26CAA0 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 239;

Db 47 AKAP 51

RESULT 36  
NAPG\_ECOLI STANDARD: PRT: 231 AA.

AC P33936; P33935;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE FERREDOXIN-TYPE PROTEIN NAPG.  
GN NAPG.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN-K12 / BHB2600;  
RA Richterich P., Lakey N., Gryan G., Jaehn L., Mintz L., Robison K.,  
RA Church G.M.;  
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blather F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
CC -1- FUNCTION: INVOLVED IN ELECTRON TRANSFER.  
CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF  
CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.  
CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A  
CC FRAMESHIFT IN POSITION 145.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U00008; AAA16398.1; ALT\_FRAME.  
DR EMBL: U00008; AAA16397.1; ALT\_FRAME.  
DR HSSP: P00195; 1CLF.  
DR EcGene: EGI2064; napg.  
DR InterPro: IPR001450; -  
DR Pfam: PF00037; fer4\_2.  
DR PROSITE: PS00198; 4FE4S\_FERREDOXIN; 1.  
RV Electon transport; Iron-sulfur; 4Fe-4S.  
FT METAL 61 61 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
FT METAL 67 67 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
FT METAL 71 71 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).  
FT METAL 99 99 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
FT METAL 102 102 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
FT METAL 107 107 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
FT METAL 111 111 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).  
FT METAL 139 139 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).  
FT METAL 147 147 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).  
FT METAL 150 150 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).  
FT METAL 154 154 IRON-SULFUR 3 (4FE-4S) (BY SIMILARITY).  
FT METAL 186 186 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).  
FT METAL 189 189 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).  
FT METAL 192 192 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).  
FT METAL 196 196 IRON-SULFUR 4 (4FE-4S) (BY SIMILARITY).  
SQ SEQUENCE 231 AA; 24924 MW; 9DA55774D6A68AD5 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 231;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OQOTA 5  
|||||  
Db 35 OQOTA 39

RESULT 37  
HIE\_CHIPA STANDARD: PRT: 235 AA.

AC P40262;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE HISTONE H1E.  
OS Chironomus pallidivittatus (Midge).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;  
OC Chironomidae; Chironomidae; Chironominae; Chironomus.  
OX NCBI\_TaxID=7151;  
[1]  
RN SEQUENCE FROM N.A.  
RA Schulze E., Mislowski J.R., Nagel S., Gavenis K., Grossbach U.;  
RL Submitted (XXX-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF  
CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.  
CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L29106; AAA21714.1; -  
DR HSSP: P08287; 1GHC.  
DR InterPro: IPR001386; -  
DR Pfam: PF00538; linker\_histone; 1.  
KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.  
SQ SEQUENCE 235 AA; 24446 MW; 340BC5B9A85002AC CRC64;

Query Match 45.5%; Score 5; DB 1; Length 235;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 AKAP 9  
|||||  
Db 28 AKAP 32

RESULT 38  
H1\_WHEAT STANDARD: PRT: 236 AA.

AC P27806;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE HISTONE H1.  
OS Triticum aestivum (Wheat).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae; Poaceae;  
OC Triticum.  
OX NCBI\_TaxID=4565;  
[1]  
RN SEQUENCE FROM N.A.  
RA STRAIN-CV. HOROSHIRIKOMUGI; TISSUE=Seedling;

```

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE ADENYLATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z47767; CAAB7696.1; -
DR HSSP: P05082; 2ECK.
DR InterPro: IPR000850; -
DR Pfam: PF00406; adenylatekinase. 1.
DR PRINTS: PR00094; ADENYLTKINASE.
DR PROSITE: PS00113; ADENYLATE_KINASE; FALSE_NEG.
DR Transferase: Kinase; ATP-binding.
FT NP_BIND 7 15 ATP (BY SIMILARITY).
SQ SEQUENCE 214 AA; 23755 MW; 55030EA82BDDAC7 CRC64;

OY 2 QOTAP 6
Db 173 QOTAP 177

Query Match 45.5%; Score 5; DB 1; Length 214;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 34
ID KAD_YERPE STANDARD; PRT; 214 AA.
AC 069172;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
GN AKK.
OS Versinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RA Munier-Lehmann H.;
RT "Adenylate kinase from Versinia pestis and Escherichia coli:
RT structural relatedness and differences linked to growth properties of
RT these two organisms.";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS SMALL UNICITOUS ENZYME IS ESSENTIAL FOR
CC MAINTENANCE AND CELL GROWTH.
CC -1- CATALYTIC ACTIVITY: ATP + AMP -> ADP + ADP.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE ADENYLATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF065382; AAC17436.1; -
DR InterPro: IPR000850; -
DR Pfam: PF00406; adenylatekinase. 1.
DR PRINTS: PR00094; ADENYLTKINASE.
DR PROSITE: PS00113; ADENYLATE_KINASE; 1.
DR Transferase: Kinase; ATP-binding.
FT NP_BIND 7 15 ATP (BY SIMILARITY).
SQ SEQUENCE 214 AA; 23755 MW; 55030EA82BDDAC7 CRC64;

```

```

SQ SEQUENCE 214 AA; 23672 MW; 36D9FD13558E86CA CRC64;

OY 2 QOTAP 6
Db 173 QOTAP 177

Query Match 45.5%; Score 5; DB 1; Length 214;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 35
ID TONB_XANCP STANDARD; PRT; 223 AA.
AC 034261;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TONB PROTEIN.
GN TONB.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98037510; PubMed=9371459;
RA Wiggerich H.G., Klaue B., Koepflin R., Pfeifer U.B., Pehler A.;
RT "Unusual structure of the tonB-exb DNA region of Xanthomonas
RT campestris pv. campestris: tonB, exbB, and exbD1 are essential for
RT ferric iron uptake, but exbD2 is not.";
RL J. Bacteriol. 179:7103-7110(1997).
CC -1- FUNCTION: INTERACTS WITH OUTER MEMBRANE RECEPTOR PROTEINS THAT
CC CARRY OUT HIGH-AFFINITY BINDING AND ENERGY DEPENDENT UPTAKE INTO
CC THE PERIPLASMIC SPACE OF SPECIFIC SUBSTRATES. IT COULD ACT TO
CC TRANSDUCE ENERGY FROM THE CYTOPLASMIC MEMBRANE TO SPECIFIC ENERGY-
CC REQUIRING PROCESSES IN THE OUTER MEMBRANE, RESULTING IN THE
CC RELEASE INTO THE PERIPLASM OF LIGANDS BOUND BY THESE OUTER
CC MEMBRANE PROTEINS (BY SIMILARITY).
CC -1- SUBUNIT: THE ACCESSORY PROTEINS EXB AND EXB2 SEEM TO FORM A
CC COMPLEX WITH TONB (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC. ANCHORED TO THE CYTOPLASMIC
CC MEMBRANE VIA ITS N-TERMINAL SIGNAL-LIKE SEQUENCE, SPANS THE
CC PERIPLASM (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TONB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z55386; CAB08610.1; -
DR Transport: Protein transport; Inner membrane; Periplasmic;
DR Transmembrane: Signal-anchor; Repeat.
FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).
FT TRANSHEM 24 44 SIGNAL-ANCHOR (POTENTIAL).
FT DOMAIN 45 223 PERIPLASMIC (POTENTIAL).
FT DOMAIN 66 78 POLY-PRO.
FT DOMAIN 82 85 POLY-PRO.
FT DOMAIN 94 101 POLY-PRO.
SQ SEQUENCE 223 AA; 23598 MW; D7C6886DF06FDA9 CRC64;

OY 5 APKAP 9
Db 173 QOTAP 177

Query Match 45.5%; Score 5; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 31
KAPS_YEAST
ID Y319_MYCPN STANDARD: PRT: 193 AA.
AC P75329;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHEETICAL. PROTEIN MG319 HOMOLOG (H08_ORF193).
GN MP454 OR MP387.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmatelaceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreith R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000037; AAB96035.1; -
KM Hypothetical protein; Transmembrane.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 136 156
SQ SEQUENCE 193 AA; 21443 MW; BAE6806C6D80C1D CRC64;

Query Match 45.5%; Score 5; DB 1; Length 193;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8
DB 68 TAPKA 72

RESULT 32
KAPS_YEAST
ID KAPS_YEAST STANDARD: PRT: 202 AA.
AC Q02196;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ADENYLYLSULFATE KINASE (EC 2.7.1.25) (APS KINASE) (ADENOSINE-
DE 5'-PHOSPHOSULFATE KINASE) (ATP ADENOSINE-5'-PHOSPHOSULFATE 3'-
DE PHOSPHOTRANSFERASE).
GN MET14 OR YKL001C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB320;
RX MEDLINE=91375456; PubMed=1654509;
RA Koroch C., Mountain H.A., Bystrom A.S.;
RT "Cloning, nucleotide sequence, and regulation of MET14, the gene
RT encoding the APS kinase of Saccharomyces cerevisiae.";
RL Mol. Gen. Genet. 229:96-108(1991).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=93070612; PubMed=1441752;
RA Duesterhoeft A., Philippsen P.;
RT "DNA sequencing and analysis of a 24.7 kb segment encompassing
RT centromere CEN11 of Saccharomyces cerevisiae reveals nine previously
RT unknown open reading frames.";
RL Yeast 8:749-759(1992).
CC CC
CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -1- CATALYTIC ACTIVITY: ATP + ADENYLYLSULFATE -> ADP +
CC 3'-PHOSPHOADENYLYLSULFATE.
CC -1- PATHWAY: SECOND STEP IN THE SULFATE ASSIMILATION PATHWAY THAT
CC LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.
CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57990; CAA41055.1; -
DR EMBL; S55315; AAB19854.1; -
DR EMBL; X65124; CAA46252.1; -
DR EMBL; Z28001; CAA81833.1; -
DR PIR; S17244; S17244.
DR SGD; S0001484; MET14.
DR InterPro; IPR002891; -
DR Pfam; PF01583; APS_kinase; 1.
KW Methionine biosynthesis; Cysteine biosynthesis; Transferase;
KW kinase; ATP-binding; Phosphorylation.
FT NP_BIND 31 38 ATP (BY SIMILARITY).
FT ACT_SITE 105 105 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 202 AA; 23060 MW; 32F301FB6B2F41D3 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 202;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9
DB 166 APKAP 170

RESULT 33
KAD_YEREN
ID KAD_YEREN STANDARD: PRT: 214 AA.
AC P43412;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ADENYLYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
GN ADR.
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6471/76 / SEROTYPE O:3;
RX MEDLINE=96100456; PubMed=8559076;
RX Skurnik M., Venho R., Toivanen P., Al-Hendy A.;
RT "A novel locus of Yersinia enterocolitica serotype O:3 involved in
RT lipopolysaccharide outer core biosynthesis.";
RL Mol. Microbiol. 17:575-594(1995).
CC CC
CC -1- FUNCTION: THIS SMALL UBQUITOUS ENZYME IS ESSENTIAL FOR
CC MAINTENANCE AND CELL GROWTH.
CC -1- CATALYTIC ACTIVITY: ATP + AMP -> ADP + ADP.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

```



DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE TRANSLATION INITIATION FACTOR IF-3.  
GN INFC.  
OS Pseudomonas fluorescens.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OX Pseudomonas.  
[1] NCBI\_TaxID=294;  
RP SEQUENCE FROM N.A.  
RC STRAIN=CHAO.  
RA Blumer C., Heeb S., Pessl G., Haas D.;  
RT Global Gaca-steered control of secondary metabolism in Pseudomonas  
RT fluorescens acts on specific ribosome binding sites."  
RL Submitted (MAR-1999) to the EMBL/Genbank/DDJ databases.  
CC -1- FUNCTION: IF-3 BINDS TO THE 30S RIBOSOMAL SUBUNIT AND SHIFTS THE  
CC EQUILIBRIUM BETWEEN 70S RIBOSOMES AND THEIR 50S AND 30S SUBUNITS IN  
CC FAVOR OF THE FREE SUBUNITS, THUS ENHANCING THE AVAILABILITY OF 30S  
CC SUBUNITS ON WHICH PROTEIN SYNTHESIS INITIATION BEGINS.  
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE IF-3 FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; AF136400; AAD34787.1; -  
DR InterPro: IPR001288; -  
DR Pfam: PF00707; IF3; 1.  
DR PROSITE: PS00938; IF3; 1.  
KM Initiation factor; Protein biosynthesis.  
SQ SEQUENCE 183 AA; 20835 MW; 003E1B9155C066A CRC64;  
  
Query Match 45.5%; Score 5; DB 1; Length 183;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 APRAP 9  
| | | | |  
DB 14 APRAP 18  
  
RESULT 29  
ID IF3\_PSESY STANDARD; PRT; 183 AA.  
AC P52834;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE TRANSLATION INITIATION FACTOR IF-3.  
GN INFC.  
OS Pseudomonas syringae (pv. syringae).  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=321;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=SUP27;  
RX MEDLINE=96198155; PubMed=8626280;  
RA Kitten T., Willis D.K.;  
RT "Suppression of a sensor kinase-dependent phenotype in Pseudomonas  
RT syringae by ribosomal proteins L35 and L20."  
RL J. Bacteriol. 178:1548-1555(1996).  
CC -1- FUNCTION: IF-3 BINDS TO THE 30S RIBOSOMAL SUBUNIT AND SHIFTS THE  
CC EQUILIBRIUM BETWEEN 70S RIBOSOMES AND THEIR 50S AND 30S SUBUNITS IN  
CC FAVOR OF THE FREE SUBUNITS, THUS ENHANCING THE AVAILABILITY OF 30S  
CC SUBUNITS ON WHICH PROTEIN SYNTHESIS INITIATION BEGINS.  
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE IF-3 FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U44118; AAB05014.1; -  
DR HSSP; P02999; 2IFE.  
DR InterPro: IPR001288; -  
DR Pfam: PF00707; IF3; 1.  
DR PROSITE: PS00938; IF3; 1.  
KM Initiation factor; Protein biosynthesis.  
SQ SEQUENCE 183 AA; 20806 MW; 50ADD828F04831FE CRC64;  
  
Query Match 45.5%; Score 5; DB 1; Length 183;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 APRAP 9  
| | | | |  
DB 14 APRAP 18  
  
RESULT 30  
ID H12\_CAEEL STANDARD; PRT; 190 AA.  
AC P15796;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE HISTONE H1.2.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN (1)  
RP SEQUENCE.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=90165873; PubMed=2407235;  
RA vanfleteren J.R., van Bun S.M., de Baere I., van Beeumen J.J.;  
RT "The primary structure of a minor isoform (H1.2) of histone H1 from  
RT the nematode Caenorhabditis elegans."  
RL Biochem. J. 265:739-746(1990).  
CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF  
CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- MISCELLANEOUS: THERE ARE 2 FORMS OF H1 IN THIS NEMATODE: H1.2 IS  
CC THE MINOR FORM.  
CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.  
CC PIR: S08195; S08195.  
DR HSSP: P08287; 1GHC.  
DR InterPro: IPR001386; -  
DR Pfam: PF00358; linker\_histone; 1.  
KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family;  
KW Acetylation.  
FT MOD\_RES 1 1 ACETYLATION (PROBABLE).  
FT DOMAIN 36 112 GLOBULAR.  
SQ SEQUENCE 190 AA; 19830 MW; FC0A04429B14C7DD CRC64;  
  
Query Match 45.5%; Score 5; DB 1; Length 190;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 APRAP 9  
| | | | |  
DB 17 APRAP 21

```

CC -----
DR EMBL: L33457; AAA67370.1; -.
DR EMBL: U50904; AAA93483.1; -.
DR HSSP: P02259; 1HST.
DR DICTYDB: DD01080; hsta.
DR InterPro: IPR001386; -.
DR Pfam: PF00538; linker_histone_1.
KW Chromosomal protein; Nuclear protein; DNA-binding; Phosphorylation.
FT INIT_MET 0
FT CONFLICT 140 179
FT SEQUENCE 179 AA: 18729 MW; ESC4528BAE3DCCFC CRC64;
SQ
Query Match
Best Local Similarity 45.5%; Score 5; DB 1; Length 179;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 6 PKAPT 10
DB 2 PKAPT 6
RESULT 26
YPOL_THEFL STANDARD: PRT; 179 AA.
ID YPOL_THEFL
AC P32438;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHELTICAL PROTEIN IN POL 5' REGION (FRAGMENT).
DE Thermus aquaticus (subsp. flavus).
CC Bacteria; Thermus/Delnococcus group; Thermus group; Thermus.
CC NCBI_TaxId=273;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ACM B-1257;
RX MEDLINE=93087201; PubMed=1454544;
RA Akmetzjanov A.A., Vakhitov V.A.;
RT "Molecular cloning and nucleotide sequence of the DNA polymerase gene
RT from Thermus flavus."
RT Nucleic Acids Res. 20:5839-5839(1992).
RL [2]
RN IDENTIFICATION, AND SIMILARITY.
RP MEDLINE=94147966; PubMed=8313894;
RX Koonin E.V., Bork P., Sander C.;
RA "Yeast chromosome III: new gene functions."
RL EMO J. 13:493-503(1994).
CC -1 SIMILARITY: BELONGS TO THE UPF0029 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X66105; -, NOT_ANNOTATED_CDS.
DR InterPro: IPR001498; -.
DR PROSITE: PS00910; UPF0029; 1.
KW Hypothetical protein.
FT NON_TER 1
FT SEQUENCE 179 AA: 19362 MW; 1B02ED2403D39786 CRC64;
SQ
Query Match
Best Local Similarity 45.5%; Score 5; DB 1; Length 179;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 5 APKAP 9
DB 105 APKAP 109

```

```

RESULT 27
NOO9_THETH STANDARD: PRT; 182 AA.
ID NOO9_THETH
AC Q56224;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 9 (EC 1.6.5.3) (NADH
DE DEHYDROGENASE 1, CHAIN 9) (NDH-1, CHAIN 9).
GN NOO9.
OS Thermus aquaticus (subsp. thermophilus).
OC Bacteria; Thermus/Delnococcus group; Thermus group; Thermus.
OX NCBI_TaxId=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HB8 / ATCC 27634;
RX MEDLINE=97172490; PubMed=9020134;
RA Yano T., Chu S.S., Sled V.D., Ohnishi T., Yagi T.;
RT "The proton-translocating NADH-quinone oxidoreductase (NDH-1) of
RT thermophilic bacterium Thermus thermophilus HB-8. Complete DNA
RT sequence of the gene cluster and thermostable properties of the
RT expressed NOO2 subunit."
RT J. Biol. Chem. 272:4201-4211(1997).
RL -1 CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1 COPFACTOR: MAY BIND TWO 4FE-4S CLUSTERS.
CC -1 SUBUNIT: COMPOSED OF 14 DIFFERENT SUBUNITS.
CC -1 SIMILARITY: BELONGS TO THE COMPLEX I 23 KDA SUBUNIT FAMILY.
CC -1 SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U52917; AAA97946.1; -.
DR HSSP: P00198; 1FDN.
DR InterPro: IPR001450; -.
DR Pfam: PF00037; fer4; 1.
DR PRINTS: PR00353; 4FE4SFERDOXIN.
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 2.
KW Oxidoreductase; NAD; Ubiquinone; Iron-sulfur; 4Fe-4S.
FT METAL 53
FT METAL 53
FT METAL 56
FT METAL 59
FT METAL 59
FT METAL 63
FT METAL 63
FT METAL 98
FT METAL 98
FT METAL 101
FT METAL 101
FT METAL 104
FT METAL 104
FT METAL 108
FT METAL 108
FT SEQUENCE 182 AA; 20080 MW; 054A9E7942C64C66 CRC64;
SQ
Query Match
Best Local Similarity 45.5%; Score 5; DB 1; Length 182;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 KAPTE 11
DB 174 KAPTE 178
RESULT 28
IF3_PSEFL STANDARD: PRT; 183 AA.
ID IF3_PSEFL
AC Q9X6E7;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

```

RP SEQUENCE FROM N.A.  
 RA MEDLINE=84011052; PubMed=6413701;  
 RA Costanzo M., Hannett N., Brzustowicz L., Pero J.;  
 RT "Bacteriophage SP01 gene 27: location and nucleotide sequence.";  
 RL J. Virol. 48:555-560(1983).  
 RN [2]  
 RP SEQUENCE OF 1-11 FROM N.A.  
 RX MEDLINE=83144054; PubMed=6402778;  
 RA Costanzo M., Pero J.;  
 RT "Structure of a Bacillus subtilis bacteriophage SP01 gene encoding  
 RT RNA polymerase sigma factor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:1236-1240(1983).  
 CC -1- FUNCTION: REQUIRED FOR LATE GENE TRANSCRIPTION AND DNA  
 CC REPLICATION.  
 CC -1- SIMILARITY: TO PHAGE SP01 GENES 44 AND 51 PROTEINS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; K01137; AAA32598.1; -;  
 DR EMBL; V01375; CAA24663.1; -;  
 KM Transcription regulation; DNA replication.  
 KW SQ SEQUENCE 155 AA; 17538 MW; 67B008ACB48C3A3 CRC64;  
 SO

Query Match 45.5%; Score 5; DB 1; Length 155;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TAPKA 8  
 I I I I I  
 DB 86 TAPKA 90

RESULT 24  
 ID FLIN\_PSEAE STANDARD; PRT; 157 AA.  
 AC Q51466;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE FLAGELLAR MOTOR SWITCH PROTEIN FLIN.  
 GN FLIN OR PA1444.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PAK;  
 RX MEDLINE=95347807; PubMed=7622217;  
 RA Simpson D.A., Ramphal R., Lory S.;  
 RT "Characterization of Pseudomonas aeruginosa fllo, a gene involved in  
 RT flagellar biosynthesis and adherence.";  
 RL Infect. Immun. 63:2950-2957(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Reiter K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Salter J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).

CC -1- FUNCTION: FLIN IS ONE OF THREE PROTEINS (FLIC, FLIN, FLIM) THAT  
 CC FORM A SWITCH COMPLEX THAT IS PROPOSED TO BE LOCATED AT THE BASE  
 CC OF THE BASAL BODY. THIS COMPLEX INTERACTS WITH THE CHEY AND CHEZ  
 CC CHEMOTAXIS PROTEINS, IN ADDITION TO CONTACTING COMPONENTS OF THE  
 CC MOTOR THAT DETERMINE THE DIRECTION OF FLAGELLAR ROTATION (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-ASSOCIATED (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE FLIN/MOP/SPAO FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L39832; AAA79754.1; -;  
 DR EMBL; AE004574; AAG04833.1; -;  
 DR InterPro; IPR001172; -;  
 DR Pfam; PF01052; Spoa; 1.  
 DR PRINTS; PR00956; FLGMOTORFLIN.  
 KM Chemotaxis; Flagella; Flagellar rotation; Inner membrane.  
 FT CONFLICT 48 P -> S (IN REF. 1).  
 SQ SEQUENCE 157 AA; 16620 MW; B3D91C0182ACB775 CRC64;  
 SO

Query Match 45.5%; Score 5; DB 1; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PKAPT 10  
 I I I I I  
 DB 62 PKAPT 66

RESULT 25  
 ID H1\_DICDI STANDARD; PRT; 179 AA.  
 AC P54671;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HISTONE H1.  
 GN H5TA.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 1-16.  
 RX MEDLINE=95172391; PubMed=7867938;  
 RA Hauser L.J., Pham M.S., Oline D.E.;  
 RT "Dictyostelium discoideum contains a single-copy gene encoding a  
 RT unique subtype of histone H1.";  
 RL Gene 154:119-122(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JH10;  
 RA O'Day D.H., Mackay L., Lydan M.A.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: HISTONES H1 ARE NECESSARY FOR THE CONDENSATION OF  
 CC NUCLEOSOME CHAINS INTO HIGHER ORDER STRUCTURES.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- PFM: PHOSPHORYLATED BY H1 HISTONE KINASE (HK).  
 CC -1- SIMILARITY: BELONGS TO THE HISTONE H1/H5 FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

RN [1]
RP SEQUENCE FROM N.A.
RA Cavan J.F., Darlous V.A., Davies C.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE UPF0076 (UK114) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U50749; AAB48552.1; -
DR InterPro: IPR000543; -
DR Pfam: PF01042; UPF0076; 1.
DR PROSITE: PS01094; UPF0076; 1.
KW Hypothetical protein.
SQ SEQUENCE 130 AA; 14232 MW; 74531BAD1F2CA525 CRC64;

```

```

Query Match 45.5%; Score 5; DB 1; Length 130;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 4 TAPKA 8
    11111
DB 10 TAPKA 14

```

```

RESULT 21
YT47 CAEEL STANDARD; PRT: 132 AA.
ID YT47 CAEEL
AC Q11074;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHEICAL 14.8 KDA PROTEIN B0416.7 IN CHROMOSOME X.
CN B0416.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-BRISTOL N2;
RL Favallo T.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: SOME, TO C.ELEGANS F36A4.2 AND F36A4.4.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U23516; AAG38883.1; -
DR WormPep: B0416.7; CEO2437.
KW Hypothetical protein.
SQ SEQUENCE 132 AA; 14811 MW; 16C3FC7580732E40 CRC64;

```

```

Query Match 45.5%; Score 5; DB 1; Length 132;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 6 PKAPT 10
    11111
DB 111 PKAPT 115

```

```

RESULT 22
DH1 HORVU STANDARD; PRT: 139 AA.
ID DH1 HORVU
AC P1251;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE DEHYDRIN DHN1 (B8).
GN DHN1.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae;
OC Hordeum.
OX NCBI_Taxid=4513;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV, HIMALAYA; TISSUE=Seedling;
RX MEDLINE=93357436; PubMed=2562763;
RA Close T.J., Kortt A.A., Chandler P.M.;
RT "A cDNA-based comparison of dehydration-induced proteins (dehydrins)
RT in barley and corn."
RL Plant Mol. Biol. 13:95-108(1989).
CC -1- INDUCTION: BY ABSCISIC ACID AND WATER-STRESS.
CC -1- SIMILARITY: VERY SIMILAR TO DHN3 AND DHN4, AND
CC ESPECIALLY TO BARLEY DHN2. TWO EXTREMELY CONSERVED BLOCKS WERE
CC IDENTIFIED, WITH A LESS CONSERVED REPEATING UNIT POSITIONED
CC BETWEEN THEM. DEHYDRIN DHN1 HAS ONLY ONE SUCH SEMI-CONSERVED
CC TANDEM REPEATING UNITS.
CC -----
CC -1- SIMILARITY: BELONGS TO THE PLANT DEHYDRIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X15288; CAA3362.1; -
DR PIR: S05544; S05544.
DR InterPro: IPR000167; -
DR Pfam: PF00257; dehydrin.1.
DR PROSITE: PS00315; DEHYDRIN.1; 1.
DR PROSITE: PS00823; DEHYDRIN.2; 2.
KW dehydrin.
FT DOMAIN 61
FT POLY-SER 69
SQ SEQUENCE 139 AA; 14236 MW; BCE6A93923BFCB CRC64;

```

```

Query Match 45.5%; Score 5; DB 1; Length 139;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 QOQTA 5
    11111
DB 102 QOQTA 106

```

```

RESULT 23
GP27 BPSP1 STANDARD; PRT: 155 AA.
ID GP27 BPSP1
AC P06228; Q38590;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENE 27 PROTEIN.
GN GENE 27.
OS Bacteriophage SP01.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae;
OC SP01-like viruses.
OX NCBI_Taxid=10685;
RN [1]

```

FT TURN 79 80  
 FT TURN 84 85  
 FT STRAND 86 86  
 FT HELIX 87 92  
 FT HELIX 101 109  
 FT TURN 110 111  
 FT HELIX 113 120  
 SO SEQUENCE 129 AA; 13976 MW; 78C2A85F3B18441 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APRAP 9  
 Db 23 APRAP 27

RESULT 19  
 ID CYC3\_DESYM STANDARD; PRT; 130 AA.  
 AC P00132; 046607;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CYTOCHROME C3 PRECURSOR.  
 OS Desulfovibrio vulgaris (strain Miyazaki).  
 OC Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.  
 NCBI\_TaxID=883;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kitamura M., Ozawa K., Kojima S., Kumagai I., Akutsu H., Miura K.I.;  
 RT "The primary structure of pre-cytochrome c(3) from Desulfovibrio  
 vulgaris (Miyazaki F) as determined by nucleotide sequencing."  
 RT gene and partial amino acid sequencing."  
 RL Protein Seq. Data Anal. 5:193-196(1993).  
 RN [2]  
 RP SEQUENCE OF 24-130.  
 RX MEDLINE=80249474; PubMed=6249799;  
 RA Shinkai W., Hase T., Yagi T., Matsubara H.;  
 RT "Amino acid sequence of cytochrome c3 from Desulfovibrio vulgaris,  
 Miyazaki.";  
 RL J. Biochem. 87:1747-1756(1980).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
 RX MEDLINE=82007753; PubMed=6268619;  
 RA Higuchi Y., Bando S., Kusunoki M., Matsuura Y., Yasuoka N., Kakudo M.,  
 RA Yamanaka T., Yagi T., Inokuchi H.;  
 RT "The structure of cytochrome c3 from Desulfovibrio vulgaris Miyazaki  
 at 2.5-A resolution."  
 RL J. Biochem. 89:1659-1662(1991).  
 RN [4]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=93044358; PubMed=1668723;  
 RA Park J.-S., Kano K., Morimoto Y., Higuchi Y., Yasuoka N., Ogata M.,  
 RA Niki K., Akutsu H.;  
 RT "H NMR studies on ferricytochrome c3 from Desulfovibrio vulgaris  
 Miyazaki F and its interaction with ferredoxin I."  
 RL J. Biomol. NMR 1:271-282(1991).  
 CC -1- FUNCTION: PARTICIPATES IN SULFATE RESPIRATION COUPLED WITH  
 PHOSPHORYLATION BY TRANSFERRING ELECTRONS FROM THE ENZYME  
 DEHYDROGENASE TO FERREDOXIN.  
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.  
 CC -1- PPM: BINDS 4 HEMES.  
 CC -1- SIMILARITY: TO OTHER C3-TYPE CYTOCHROMES.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch).

CC EMBL: D31702; BAA06511.1; -.  
 CC PIR: A00125; CCDV3M.  
 DR PIR; 2CDV; 21-FEB-84.  
 DR InterPro: IPR000345; -.  
 DR InterPro: IPR002322; -.  
 DR Pfam: PF02085; Cytochrome CIII; 1.  
 DR PRINTS: PR00609; CYTOCHROME C3.  
 DR PROSITE: PS00190; CYTOCHROME C; 2.  
 KW Electron transport; Sulfate respiration; Heme; Periplasmic; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 130  
 FT METAL 45 45  
 FT METAL 48 48  
 FT METAL 53 53  
 FT BINDING 56 56  
 FT METAL 57 57  
 FT METAL 58 58  
 FT BINDING 69 69  
 FT BINDING 74 74  
 FT METAL 75 75  
 FT METAL 93 93  
 FT BINDING 102 102  
 FT BINDING 105 105  
 FT METAL 106 106  
 FT BINDING 123 123  
 FT BINDING 128 128  
 FT METAL 129 129  
 FT METAL 65 65  
 FT STRAND 32 34  
 FT STRAND 41 43  
 FT HELIX 46 48  
 FT TURN 49 50  
 FT HELIX 53 56  
 FT STRAND 59 60  
 FT TURN 61 62  
 FT STRAND 63 64  
 FT TURN 69 70  
 FT TURN 72 73  
 FT STRAND 77 77  
 FT TURN 80 81  
 FT TURN 85 86  
 FT STRAND 87 87  
 FT HELIX 88 93  
 FT HELIX 102 110  
 FT TURN 111 112  
 FT HELIX 114 121  
 SO SEQUENCE 130 AA; 13843 MW; CCBBA60351B9EF8E CRC64;

Query Match 45.5%; Score 5; DB 1; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 31;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APRAP 9  
 Db 24 APRAP 28

RESULT 20  
 ID Y142\_LEUWC STANDARD; PRT; 130 AA.  
 AC P97117;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HYPOTHETICAL 14.2 KDA PROTEIN.  
 OS Leuconostoc mesenteroides (subsp. cremoris).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
 OC Leuconostoc.  
 NCBI\_TaxID=33965;

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: U37245; AAC22286.1; -  
 DR TIGR: H10626; -  
 DR InterPro: IPR001185; -  
 DR Pfam: PF01741; MSCL: 1.  
 DR PRINTS: PRO1264; MECHCHANNEL.  
 DR PROSITE: PS01327; MSCL: 1.  
 KW Transmembrane; Inner membrane; Ionic channel.  
 FT DOMAIN 1 15 CYTOPLASMIC (BY SIMILARITY).  
 FT TRANSMEM 16 42 BY SIMILARITY.  
 FT DOMAIN 43 74 PERIPLASMIC (BY SIMILARITY).  
 FT TRANSMEM 75 96 BY SIMILARITY.  
 FT DOMAIN 97 128 CYTOPLASMIC (BY SIMILARITY).  
 FT SEQUENCE 128 AA: 14198 MW: 803A68D27200E51A CRC64.  
 SO  
 Query Match 45.5%; Score 5; DB 1; Length 128;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 4 TAPKA 8  
 11111  
 Db 110 TAPKA 114  
 RESULT 18  
 CYC3\_DESVH STANDARD: PRT: 129 AA.  
 ID P00131;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CYTOCHROME C3 PRECURSOR.  
 OS Desulfovibrio vulgaris (strain Hildenborough).  
 OC Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.  
 OX NCBI\_Taxid=882;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87004646; PubMed=3019687;  
 RA Voordouw G., Brenner S.;  
 RT "Cloning and sequencing of the gene encoding cytochrome c3 from  
 RT Desulfovibrio vulgaris (Hildenborough).";  
 RL Eur. J. Biochem. 159:347-351(1986).  
 RN [2]  
 RP SEQUENCE OF 23-129.  
 RX STRAIN-NCIB 8303;  
 RC MEDLINE=74070664; PubMed=4358550;  
 RA Troussil E.B., Campbell L.L.;  
 RT "Amino acid sequence of cytochrome c3 from Desulfovibrio vulgaris,";  
 RL J. Biol. Chem. 249:386-393(1974).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=92138655; PubMed=1663945;  
 RA Morimoto Y., Tanl T., Okumura H., Higuchi Y., Yasuoka N.;  
 RT "Effects of amino acid substitution on three-dimensional structure:  
 RT an x-ray analysis of cytochrome c3 from Desulfovibrio vulgaris  
 RT Hildenborough at 2-A resolution.";  
 RL J. Biochem. 110:532-540(1991).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
 RX MEDLINE=94076344; PubMed=8254667;  
 RA Matias P.M., Piraao C., Morais J., Coll M., Carrondo M.A.;  
 RT "Structure analysis of cytochrome c3 from Desulfovibrio vulgaris  
 RT Hildenborough at 1.9-A resolution.";  
 RL J. Mol. Biol. 234:680-699(1993).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS).  
 RX Simoes P., Matias P.M., Morais J., Wilson K., Dauter Z.,  
 RA Carrondo M.A.;  
 RL Submitted (JUN-1997) to the PDB data bank.  
 RN [6]

RP STRUCTURE BY NMR.  
 RX MEDLINE=93093162; PubMed=1333991;  
 RA Salgueiro C.A., Turner D.L., Santos H., Legall J., Xavier A.V.;  
 RT "Assignment of the redox potentials to the four hemes in  
 RT Desulfovibrio vulgaris cytochrome c3 by 2D-NMR.";  
 RL FEBS Lett. 314:155-158(1992).  
 RN [7]  
 RP STRUCTURE BY NMR.  
 RX MEDLINE=98378464; PubMed=9710542;  
 RA Mesias A.C., Kastan D.H.W., Costa H.S., Legall J., Turner D.L.,  
 RA Santos H., Xavier A.V.;  
 RT "Solution structure of desulfovibrio vulgaris (Hildenborough)  
 RT ferredoxin cytochrome c3: structural basis for functional cooperativity.";  
 RL J. Mol. Biol. 281:719-739(1998).  
 CC -1- FUNCTION: PARTICIPATES IN SULFATE RESPIRATION COUPLED WITH  
 CC PHOSPHORYLATION BY TRANSFERRING ELECTRONS FROM THE ENZYME  
 CC DEHYDROGENASE TO FERREDOXIN.  
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.  
 CC -1- PTM: BINDS 4 HEMES.  
 CC -1- SIMILARITY: TO OTHER C3-TYPE CYTOCHROMES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X04304; CAA27847.1; -  
 DR PIR: A00124; CDDV3.  
 DR PIR: A24799; A24799.  
 DR PDB: 2CTH; 24-DEC-97.  
 DR PDB: 2CYM; 30-APR-94.  
 DR PDB: 3CYR; 28-JAN-98.  
 DR PDB: 1A21; 08-JUL-98.  
 DR InterPro: IPR000345; -  
 DR InterPro: IPR002322; -  
 DR Pfam: PF02085; Cytochrome\_CIII; 1.  
 DR PRINTS: PRO0609; CYTOCHROME\_C3.  
 DR PROSITE: PS00190; CYTOCHROME\_C; 2.  
 KW Electron transport; Sulfate respiration; Heme; Periplasmic; Signal;  
 KW 3D-structure.  
 FT SIGNAL 1 22  
 FT CHAIN 23 129  
 FT METAL 34 34 CYTOCHROME C3  
 FT METAL 37 37 IRON (HEME 1 AXIAL LIGAND).  
 FT BINDING 52 52 IRON (HEME 3 AXIAL LIGAND).  
 FT BINDING 55 55 HEME 1 (COVALENT).  
 FT METAL 56 56 HEME 1 (COVALENT).  
 FT METAL 57 57 IRON (HEME 1 AXIAL LIGAND).  
 FT BINDING 68 68 IRON (HEME 2 AXIAL LIGAND).  
 FT BINDING 73 73 HEME 2 (COVALENT).  
 FT METAL 74 74 HEME 2 (COVALENT).  
 FT METAL 92 92 IRON (HEME 2 AXIAL LIGAND).  
 FT BINDING 101 101 IRON (HEME 4 AXIAL LIGAND).  
 FT BINDING 104 104 HEME 3 (COVALENT).  
 FT METAL 105 105 IRON (HEME 3 AXIAL LIGAND).  
 FT BINDING 122 122 HEME 4 (COVALENT).  
 FT BINDING 127 127 HEME 4 (COVALENT).  
 FT METAL 128 128 IRON (HEME 4 AXIAL LIGAND).  
 FT STRAND 31 33  
 FT STRAND 40 42  
 FT HELIX 45 47  
 FT TURN 48 49  
 FT HELIX 52 54  
 FT TURN 55 55  
 FT STRAND 58 59  
 FT TURN 60 61  
 FT STRAND 62 63  
 FT TURN 68 69  
 FT TURN 71 72  
 FT STRAND 76 76

DR Pfam: PF00379; Insect\_cuticle; 1.  
 DR PRINTS: PR00947; CUTICLE.  
 DR PROSITE: PS00233; CUTICLE; 1.  
 KM Structural protein; Cuticle.  
 FT DOMAIN 73  
 SQ SEQUENCE 126 AA; 12421 MW; F9633E26A877F6DE CRC64;

Query Match 45.5%; Score 5; DB 1; Length 126;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 Db 83 APKAP 87

RESULT 15  
 CU26\_ARADI STANDARD; PRT; 127 AA.  
 AC P80517;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 13-JUL-1998 (Rel. 36, Last annotation update)  
 DE ADULT-SPECIFIC RIGID CUTICULAR PROTEIN 12.6 (ACP 12.6).  
 OS Araneus diadematus (Spider).  
 CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;  
 CC Araneomorphae; Entelegynae; Araneidae; Araneidae; Araneus.  
 OX NCBI\_Taxid=45920;  
 RN [1]  
 RC SEQUENCE.  
 RP TISSUE-Cuticle;  
 RX MEDLINE=97166616; Pubmed=9014336;  
 RA Norup T., Berg T., Stenholm H., Andersen S.O., Hoefrup P.;  
 RT "Purification and characterization of five cuticular proteins from  
 the spider Araneus diadematus".  
 RL Insect Biochem. Mol. Biol. 26:907-915(1996).  
 CC -1- FUNCTION: COMPONENT OF THE RIGID CUTICLE OF THE SPIDER.  
 CC -1- MASS SPECTROMETRY: MW=12587.5; METHOD-ELECTROSPRAY.  
 CC -1- SIMILARITY: CONTAINS A CUTICLE CONSENSUS DOMAIN.  
 DR Interpro: IPR000618;  
 DR Pfam: PF00379; Insect\_cuticle; 1.  
 DR PRINTS: PR00947; CUTICLE.  
 DR PROSITE: PS00233; CUTICLE; 1.  
 KM Structural protein; Cuticle.  
 FT DOMAIN 73  
 SQ SEQUENCE 127 AA; 12589 MW; 9BFBCFB2B6AA4B9 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 Db 83 APKAP 87

RESULT 16  
 ULB0\_HCMVA STANDARD; PRT; 127 AA.  
 ID ULB0\_HCMVA  
 AC P16830;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last annotation update)  
 DE HYPOTHETICAL PROTEIN ULI10.  
 GN ULI10.  
 OS Human cytomegalovirus (strain AD169).  
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 CC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_Taxid=10360;  
 RN [1]  
 SQ SEQUENCE FROM N.A.

RX MEDLINE=90269039; Pubmed=2161319;  
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,  
 RA Horsnell T., Hutchison C.A. III, Kourzides T., Martignetti J.A.,  
 RA Predde E., Satchell S.C., Tomlinson P., Weston K.M., Barrell B.G.;  
 RT "Analysis of the protein-coding content of the sequence of human  
 cytomegalovirus strain AD169."  
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb.slb.ch](mailto:license@isb.slb.ch)).

DR EMBL: X17403; CA35348.1;  
 DR PIR: S09877; S09877.  
 DR Hypothetical protein.  
 KM SEQUENCE 127 AA; 14224 MW; D788B6E5C864F5AB CRC64;

Query Match 45.5%; Score 5; DB 1; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QOTAP 6  
 |||||  
 Db 23 QOTAP 27

RESULT 17  
 MSCL\_HAEIN STANDARD; PRT; 128 AA.  
 ID MSCL\_HAEIN  
 AC P44789;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE LARGE-CONDUCTANCE MECHANOSENSITIVE CHANNEL.  
 GN MSCL OR H10626  
 OS Haemophilus influenzae.  
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 CC Haemophilus.  
 OX NCBI\_Taxid=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; Pubmed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Ueterbach T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 influenzae Rd."  
 RL Science 269:496-512(1995).  
 CC -1- FUNCTION: CHANNEL THAT OPENS IN RESPONSE TO STRETCH FORCES IN THE  
 CC MEMBRANE LIPID BILAYER. MAY PARTICIPATE IN THE REGULATION OF  
 CC OSMOTIC PRESSURE CHANGES WITHIN THE CELL (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMOPENTAMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
 CC (BY SIMILARITY).  
 CC -----  
 CC -1- SIMILARITY: BELONGS TO THE MSCL FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb.slb.ch](mailto:license@isb.slb.ch)).

FT NON\_TER 123 123  
SQ SEQUENCE 123 AA: 14162 MW: 62E73FB457EBB51C CRC64:

Query Match  
Best Local Similarity 45.5%; Score 5; DB 1; Length 123;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 KAPTE 11  
DB 56 KAPTE 60

RESULT 12  
WN3A\_MELGA STANDARD; PRT: 123 AA.  
ID WN3A\_MELGA  
AC P28125;  
DT 01-DEC-1992 (Rel. 24, Created)  
DR 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 01-JUL-1993 (Rel. 26, Last annotation update)  
DE WNT-3A PROTEIN (FRAGMENT).  
GN Meleagris gallopavo (Common turkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.  
OX NCBI\_TaxId=9103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92279273; PubMed=1534411;  
RA Sidorow A.;  
RT "Diversification of the Wnt gene family on the ancestral lineage of vertebrates."  
RL Proc. Natl. Acad. Sci. U.S.A. 89:5098-5102(1992).  
CC -1- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.  
CC -1- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE EXTRACELLULAR MATRIX.  
CC -1- SIMILARITY: BELONGS TO THE WNT FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: M91281; AAA49630.1; -  
DR InterPro: IPR000970; -  
DR Pfam: PF00110; Wnt: 1.  
DR PROSITE: PS00246; WNT1; PARTIAL.  
KW Developmental protein; Glycoprotein.  
FT NON\_TER 1 123  
FT NON\_TER 1 123  
SQ SEQUENCE 123 AA: 14268 MW: 66B6CEFF05BBA576 CRC64:

Query Match  
Best Local Similarity 45.5%; Score 5; DB 1; Length 123;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 KAPTE 11  
DB 56 KAPTE 60

RESULT 13  
WN3A\_SCEOC STANDARD; PRT: 123 AA.  
ID WN3A\_SCEOC  
AC P28142;  
DT 01-DEC-1992 (Rel. 24, Created)  
DR 01-DEC-1992 (Rel. 24, Last sequence update)

DT 01-JUL-1993 (Rel. 26, Last annotation update)  
DE WNT-3A PROTEIN (FRAGMENT).  
GN WNT-3A.  
OS Sceloporus occidentalis (Western fence lizard).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Iguania; Iguanidae; Phrynosomatinae;  
OC Sceloporus.  
OX NCBI\_TaxId=8519;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92279273; PubMed=1534411;  
RA Sidorow A.;  
RT "Diversification of the Wnt gene family on the ancestral lineage of vertebrates."  
RL Proc. Natl. Acad. Sci. U.S.A. 89:5098-5102(1992).  
CC -1- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.  
CC -1- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE EXTRACELLULAR MATRIX.  
CC -1- SIMILARITY: BELONGS TO THE WNT FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: M91299; AAA49542.1; -  
DR InterPro: IPR000970; -  
DR Pfam: PF00110; Wnt: 1.  
DR PROSITE: PS00246; WNT1; PARTIAL.  
KW Developmental protein; Glycoprotein.  
FT NON\_TER 1 123  
FT NON\_TER 1 123  
SQ SEQUENCE 123 AA: 14296 MW: 66A1E9CF05BBBC6F CRC64:

Query Match  
Best Local Similarity 45.5%; Score 5; DB 1; Length 123;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 KAPTE 11  
DB 56 KAPTE 60

RESULT 14  
CU24\_ARADI STANDARD; PRT: 126 AA.  
ID CU24\_ARADI  
AC P80516;  
DT 01-FEB-1996 (Rel. 33, Created)  
DR 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE ADULT-SPECIFIC RIGID CUTICULAR PROTEIN 12.4 (ACP 12.4).  
OS Araneus diadematus (Spider).  
OC Araneura; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;  
OC Araneomorphae; Entelegynae; Araneidae; Araneidae; Araneus.  
OX NCBI\_TaxId=45920;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Cuticle;  
RX MEDLINE=9716616; PubMed=9014336;  
RA Norup T., Berg T., Stenholm H., Andersen S.O., Hoejrup P.;  
RT "Purification and characterization of five cuticular proteins from the spider Araneus diadematus."  
RL Insect Biochem. Mol. Biol. 26:907-915(1996).  
CC -1- FUNCTION: COMPONENT OF THE RIGID CUTICLE OF THE SPIDER.  
CC -1- MASS SPECTROMETRY: MM=12420.5; METHOD=ELECTROSPRAY.  
CC -1- SIMILARITY: CONTAINS A CUTICLE CONSENSUS DOMAIN.  
DR InterPro: IPR000618; -



FT MOD.RES 6 METHYLATION (BY SIMILARITY).  
SQ SEQUENCE 119 AA: 13339 MW: 826ECB33A26BD873 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 119;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
DB 102 TAPKA 106

RESULT 9  
CU19\_ARADI STANDARD; PRT; 120 AA.

AC P80515;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE ADULT-SPECIFIC RIGID CUTICULAR PROTEIN 11.9 (ACP 11.9).  
OS Araneus diadematus (Spider)  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;  
OC Araneomorphae; Entelegynae; Araneidae; Araneidae; Araneus.  
OX NCBI\_TaxID=45920;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Cuticle;  
RX MEDLINE=97166616; PubMed=9014336;  
RA Norup T., Berg T., Stenholm H., Andersen S.O., Hoejrup P.;  
RT "Purification and characterization of five cuticular proteins from  
the spider Araneus diadematus".  
RL Insect Biochem. Mol. Biol. 26:907-915(1996).  
CC -1- FUNCTION: COMPONENT OF THE RIGID CUTICLE OF THE SPIDER.  
CC -1- MASS SPECTROMETRY: MW=11872.1; METHOD-ELECTROSPRAY.  
CC -1- SIMILARITY: CONTAINS A CUTICLE CONSENSUS DOMAIN.  
DR HSSP: P04002; IATF.  
DR InterPro: IPR000618;  
DR Pfam: PF00379; InsectCuticle; 1.  
DR PRINTS: PR00947; CUTICLE.  
DR PROSITE: PS00233; CUTICLE; 1.  
KW Structural protein; Cuticle.  
FT DOMAIN 73 83 POLY-ALA.  
SO SEQUENCE 120 AA: 11871 MW: 8D3E577C6199DC2 CRC64;

Query Match 45.5%; Score 5; DB 1; Length 120;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
DB 83 APRAP 87

RESULT 10  
H2B\_HOLTU STANDARD; PRT; 123 AA.

AC P48557;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE HISTONE H2B.  
OS Holothuria tubulosa (Sea cucumber).  
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinzoa;  
OC Holothuroidea; Aspidochelerozoa; Aspidochelerozoa; Holothuriidae;  
OC Holothuria.  
OX NCBI\_TaxID=7685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Drabant B., Louroutziatis A., Prats E., Cornudella L., Doenecke D.;  
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBUNIT: THE NUCLEOSOME IS AN OCTAMER CONTAINING TWO MOLECULES OF

CC H2A, H2B, H3, AND H4; WHICH WRAP APPROXIMATELY 146 BP OF DNA.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- SIMILARITY: BELONGS TO THE HISTONE H2B FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Z46225; CAA86297.1; -  
DR InterPro: IPR000166; -  
DR InterPro: IPR000558; -  
DR Pfam: PF00123; Histone; 1.  
DR PRINTS: PR00621; HISTONEH2B.  
DR PROSITE: PS00357; HISTONEH2B; 1.  
SQ Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding.  
SQ SEQUENCE 123 AA: 13649 MW: 35503AFB1400FE6E CRC64;

Query Match 45.5%; Score 5; DB 1; Length 123;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
DB 2 APRAP 6

RESULT 11  
WN3A\_ALOVU STANDARD; PRT; 123 AA.

AC P28101;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE WNT-3A PROTEIN (FRAGMENT).  
GN WNT-3A.  
OS Alopias vulpinus (Thresher shark).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphi; Galeoidea; Lamniformes; Alopiidae;  
OC Alopiidae; Alopias.  
OX NCBI\_TaxID=7852;

Query Match 45.5%; Score 5; DB 1; Length 123;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
DB 2 APRAP 6

RESULT 11  
WN3A\_ALOVU STANDARD; PRT; 123 AA.

AC P28101;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE WNT-3A PROTEIN (FRAGMENT).  
GN WNT-3A.  
OS Alopias vulpinus (Thresher shark).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphi; Galeoidea; Lamniformes; Alopiidae;  
OC Alopiidae; Alopias.  
OX NCBI\_TaxID=7852;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92279273; PubMed=1534411;  
RA Sidor A.;  
RT "Diversification of the Wnt gene family on the ancestral lineage of  
vertebrates".  
RL Proc. Natl. Acad. Sci. U.S.A. 89:5098-5102(1992).  
CC -1- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING  
CC MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF  
CC TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIMETERS.  
CC -1- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE  
CC EXTRACELLULAR MATRIX.  
CC -1- SIMILARITY: BELONGS TO THE WNT FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

DE 15-DEC-99/8 (Rel. 37, last annotation update)
DT RNA REPLICATION PROTEIN (150 KDA PROTEIN) (ORF 1) [CONTAINS: RNA-
DE DIRECTED RNA POLYMERASE (EC 2.7.7.48); PROBABLE HELICASE].
OS Strawberry mild yellow edge-associated virus (SMYAV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potexvirus.
OX NCBI_TaxId=12187;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MY-18;
RX MEDLINE=92166762; PubMed=1339469;
RA Jekmann W., Mais E., Martin R.R.;
RT "The nucleotide sequence and genome organisation of strawberry mild
RL yellow edge-associated potexvirus.";
RT J. Gen. Virol. 73:475-479(1992).
CC -1- FUNCTION: RNA-REPLICATION. THE CENTRAL PART OF THIS PROTEIN
CC POSSIBLY FUNCTIONS AS A ATP-BINDING HELICASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on way
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D12517; BAA02082.1; -.
DR PIR; JQ1426; RRMQSM.
DR InterPro; IPR000606; -.
DR Pfam; PF01443; Viral_helicase1.1.
DR ATP-binding; Helicase; RNA replication; RNA-directed RNA polymerase;
RW Transferase.
KW NP_BIND 605 612 POTENTIAL.
SO SEQUENCE 1323 AA; 149593 MW; D6BC35133E3CAD26 CRC64;

```

Query Match	54.5%	Score 6:	DB 1:	Length 1323:
Best Local Similarity	100.0%	Pred. No. 18:		
Best Match	6:	Conservative	0:	Mismatches 0:
				Indels 0:
				Gaps 0:

Qy	5	APKAPT	10
Db	463	APKAPT	468

RESULT	7
ID	ANT1_ONCGVO
AC	ANT1_ONCGVO
AC	P21249;
DT	01-MAY-1991 (Rel. 18, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	MAJOR ANTIGEN.
GN	OV1.
OS	Onchocerca volvulus.
CC	Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC	Onchocercidae; Onchocerca.
OX	NCBI_TaxID=6282;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=95287898; PubMed=7770081;
RA	Triteerapapab S., Ritchie T.L., Tuan R.S., Shepley K.J., Dimman J.D.,
RA	Neubert T.A., Scott A.L.;
RT	"Molecular cloning of a gene expressed during early embryonic
RT	development in Onchocerca volvulus.";
RL	Mol. Biochem. Parasitol. 69:161-171(1995).
RL	[2]
RP	SEQUENCE OF 733-866 FROM N.A.
RX	MEDLINE=89127417; PubMed=2464764;
RA	Donelson J.E., Duke B.O.L., Moser D., Zeng W., Erondy N.E.,
RA	Lactus R., Renz A., Karam M., Flores G.Z.;
RT	"Construction of Onchocerca volvulus cDNA libraries and partial
RT	characterization of the cDNA for a major antigen.";
RL	Mol. Biochem. Parasitol. 31:241-250(1988).

```

CC      -1- FUNCTION: MAY BE A MYOFIBRILLAR PROTEIN.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC
DR      EMBL: U12681; AAA80009.1; -.
DR      EMBL: J03995; AAA29412.1; -.
DR      PIR: A54513; A54513.
DR      HSSP: P02633; 3ICB.
DR      Antigen: Coiled coil.
KW      DOMAIN 74 120 COILED COIL (POTENTIAL).
FT      DOMAIN 151 251 COILED COIL (POTENTIAL).
FT      DOMAIN 327 384 COILED COIL (POTENTIAL).
FT      DOMAIN 417 1879 COILED COIL (POTENTIAL).
SQ      SEQUENCE 2022 AA; 237341 MW; B7132ACFE1520317 CRC64;

```

Query Match	54.5%	Score 6:	DB 1:	Length 2022:
Similarity	100.0%	Pred. NO. 26:		
Best Local				
Matches	6;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

QY	1	QQQTAP	6
Db	302	QQQTAP	307

RESULT	8	
GSPI_AERHY		
ID_GSPI_AERHY		STANDARD;
		PRT;
		119 AA

DT 01-JUL-1993 (Rel. 26, last sequence update)  
DT 01-FEB-1995 (Rel. 31, last annotation update)  
DE GENERAL SECRETION PATHWAY PROTEIN I PRECURSOR

05 *Aeromonas hydrophila*.  
0C Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;  
0C *Aeromonas*.

RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-AH65;

RX MEDLINE=94012544; PubMed=84079845;  
 RA Howard S.P., Critch J., Bedi A.;  
 RT "Isolation and analysis of eight exo genes and their involvement in  
 RT extracellular protein secretion and outer membrane assembly in  
 RT *Aeromonas hydrophila*.";

CC -1- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE EXPORT OF PROTEINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X08504; CAA4130.1; -  
DR PIR; S22912; S22912.  
DR PIR; B49905; B49905.  
DR InterPro; IPR00130; -

DR	PROSITE: PS00409; PROKAR_NTER_METHYL; 1.
KW	Transport; Methylation.
FT	PROPEP 1 5 BY SIMILARITY.
FT	CHAIN 6 119 GENERAL SECRETION PATHWAY PROTEIN I



FT NON\_TER 314 314  
SQ SEQUENCE 314 AA; 34618 MM; 4F949656A8BED1F6 CRC64;

Query Match 54.5%; Score 6; DB 1; Length 314;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAPT 10  
Db 49 APRAPT 54

RESULT 3  
P53\_MARMO STANDARD; PRT: 391 AA.  
ID P53\_MARMO  
AC 036006;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE CELLULAR TUMOR ANTIGEN P53 (TUMOR SUPPRESSOR P53).  
GN TP53.  
OS Marmota monax (Woodchuck).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;  
OC Marmota.  
OX NCBI\_TaxID=9995;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97376966; PubMed=9233767;  
RA Feltelson M.A., Ranganathan P.N., Clayton M.M., Zhang S.M.;  
RT "Partial characterization of the woodchuck tumor suppressor, p53, and  
its interaction with woodchuck hepatitis virus X antigen in  
hepatocarcinogenesis";  
RT Oncogene 15:327-336(1997).  
RL -1- FUNCTION: ACTS AS A TUMOR SUPPRESSOR IN MANY TUMOR TYPES. INDUCES  
GROWTH ARREST OR APOPTOSIS DEPENDING ON THE PHYSIOLOGICAL  
CIRCUMSTANCES OR CELL TYPE, BUT BOTH ACTIVITIES ARE INVOLVED IN  
TUMOR SUPPRESSION. ACTS IN CELL CYCLE REGULATION, IS A TRANS-  
ACTIVATOR THAT ACTS TO NEGATIVELY REGULATE CELL DIVISION BY  
CONTROLLING A SET OF GENES REQUIRED FOR THIS PROCESS. ONE OF THE  
ACTIVATED GENES IS AN INHIBITOR OF CYCLIN-DEPENDENT KINASES.  
CC APOPTOSIS INDUCTION SEEMS TO BE MEDIATED EITHER BY STIMULATION OF  
BAX AND FAS ANTIGEN EXPRESSION, OR BY REPRESSION OF BCL-2  
EXPRESSION.  
CC -1- SUBUNIT: BINDS DNA AS AN HOMOTETRAMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- DISEASE: P53 IS FOUND IN INCREASED AMOUNTS IN A WIDE VARIETY  
OF TRANSFORMED CELLS. P53 IS FREQUENTLY MUTATED OR INACTIVATED  
IN MANY TYPES OF CANCER.  
CC -1- SIMILARITY: BELONGS TO THE P53 FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AJ001022; CA004478.1; -  
DR HSSP: P04637; ITR.  
DR InterPro: IPR002117; -  
DR Pfam: PF00870; P53; 1.  
DR PRINTS: PR00386; P53SUPPRESSR.  
DR PROSITE: PS00348; P53; 1.  
KW Anti-oncogene; DNA-binding; Transcription regulation; Activator;  
KW Nuclear protein; Phosphorylation.  
FT DOMAIN 1 44 TRANSCRIPTION ACTIVATION (ACIDIC).  
FT DNAS\_BIND 100 290 BY SIMILARITY.  
FT DOMAIN 323 354 OLIGOMERIZATION.  
FT DOMAIN 366 385 BASIC (REPRESSION OF DNA-BINDING).  
FT DOMAIN 309 321 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT MOD\_RES 390 390 PHOSPHORYLATION (BY SIMILARITY).  
SQ SEQUENCE 391 AA; 43468 MM; E1DE5DB84BA0182 CRC64;

Query Match 54.5%; Score 6; DB 1; Length 391;  
Best Local Similarity 100.0%; Pred. No. 6.4;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAPT 10  
Db 69 APRAPT 74

RESULT 4  
IF2\_SYNY3 STANDARD; PRT: 1001 AA.  
ID IF2\_SYNY3  
AC P72689;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE TRANSLATION INITIATION FACTOR IF-2.  
GN INP OR SLR0744.  
OS Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
OC NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nario K.,  
RA Okumura S., Shimpō S., Takeuchi C., Wada T., Watanabe A.,  
RA Yanada M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
Synecocystis sp. strain PCC6803. II. Sequence determination of the  
entire genome and assignment of potential protein-coding regions.";  
RT DNA Res. 3:109-136(1996).  
RL -1- FUNCTION: IF-2, ONE OF THE ESSENTIAL COMPONENTS FOR THE INITIATION  
OF PROTEIN SYNTHESIS IN VITRO, PROTECTS FORMYL METHIONYL-TRNA FROM  
SPONTANEOUS HYDROLYSIS AND PROMOTES ITS BINDING TO THE 30S  
RIBOSOMAL SUBUNIT. IT IS ALSO INVOLVED IN THE HYDROLYSIS OF GTP  
DURING THE FORMATION OF THE 70S RIBOSOMAL COMPLEX (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE IF-2 FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D90900; BA16696.1; -  
DR HSSP: P02990; IETU.  
DR InterPro: IPR000178; -  
DR InterPro: IPR000795; -  
DR Pfam: PF00009; GTP\_EFTU; 1.  
DR Pfam: PF02131; IF2; 1.  
DR PROSITE: PS01176; IF2; 1.  
KW Initiation factor; Protein biosynthesis; GTP-binding.  
FT DOMAIN 494 648 G-DOMAIN.  
FT NP\_BIND 502 509 GTP (BY SIMILARITY).  
FT NP\_BIND 552 556 GTP (BY SIMILARITY).  
FT NP\_BIND 606 609 GTP (BY SIMILARITY).  
SQ SEQUENCE 1001 AA; 108118 MM; 4C7F00DEB56CE3C CRC64;

Query Match 54.5%; Score 6; DB 1; Length 1001;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAPT 10

```

983 4 36.4 437 1 E2F2_HUMAN 014209 homo sapien
984 4 36.4 437 1 SAHH_CABEL P27604 caenorhabdi
985 4 36.4 438 1 FUMC_SULSO P39461 scutolobus
986 4 36.4 438 1 PRIL_DROME 024317 drosophila
987 4 36.4 438 1 STD_PYRRO 058776 pyrococcus
988 4 36.4 438 1 STD_PYRRO 052428 pyrococcus
989 4 36.4 438 1 Y2B1_HAEIN P44610 haemophilus
990 4 36.4 439 1 CK13_SCHPO 074135 schizosacch
991 4 36.4 439 1 RHG1_HUMAN 007960 homo sapien
992 4 36.4 440 1 MANA_CANAL P34948 candida alb
993 4 36.4 441 1 CG22_ANTMA P34801 anltrithum
994 4 36.4 442 1 Y141_ECOLI P03835 escherichia
995 4 36.4 443 1 AROA_BACNO P04650 bacteroides
996 4 36.4 443 1 CAR4_DICDI 091430 dictyostel
997 4 36.4 443 1 DCUA_HELPI 092100 helicobacte
998 4 36.4 443 1 DCUA_HELPI 025425 helicobacte
999 4 36.4 443 1 SOX3_HUMAN P41225 homo sapien
1000 4 36.4 443 1 YJ01_ECOLI P39410 escherichia

```

## ALIGNMENTS

```

RESULT 1
ID P60_LISMO STANDARD: PRT: 484 AA.
AC P21171:
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROTEIN P60 PRECURSOR (INVASION-ASSOCIATED PROTEIN).
GN IAP.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-49.
RC STRAIN-SV1/2A EGD(KAUFMANN);
RX MEDLINE=90256283; PubMed=2111287;
RA Koehler S., Lelmeister-Waechter M., Chakraborty T., Lottspeich F.,
RA Goebel W.;
RT "The gene coding for protein p60 of Listeria monocytogenes and its
RT use as a specific probe for Listeria monocytogenes.";
RN Infect. Immun. 58:1943-1950(1990).
RL [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=93094153; PubMed=1459966;
RA Buber A., Kuhn M., Goebel W., Koehler S.;
RT "Structural and functional properties of the p60 proteins from
RT different Listeria species.";
RL J. Bacteriol. 174:8166-8171(1992).
CC -1- FUNCTION: THIS MAJOR EXTRACELLULAR PROTEIN MAY BE INVOLVED IN THE
CC INVASION OF NONPROFESSIONAL PHAGOCYTTIC CELLS BY LISTERIA.
CC -1- SIMILARITY: BELONGS TO THE E.COLI NLPC / LISTERIA P60 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X53268; CA36509.1; -.
CC PIR: A41487; A41487.
CC InterPro: IPR000064; -.
CC DR InterPro: IPR002482; -.
CC Pfam: PF00877; NLPC_P60; 1.
CC KMW Pfam: PF01476; PG_binding_2; 2.
CC Signal: Repeat.
CC FT SIGNAL 1 27
CC FT CHAIN 28 484 PROTEIN P60.

```

```

FT DOMAIN 311 355 19 X 2 AA TANDEM REPEATS OF T-N.
SQ SEQUENCE 484 AA; 50587 MW; 3CC0F90591E14E0F CRC64;

```

```

Query Match 100.0%; Score 11; DB 1; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QOOTAPKAPTE 11
Db 291 QOOTAPKAPTE 301

```

```

RESULT 2
ID P53_SPEBE STANDARD: PRT: 314 AA.
AC 064662;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CELLULAR TUMOR ANTIGEN P53 (TUMOR SUPPRESSOR P53) (FRAGMENT).
GN TP53.
OS Sperophilus becheylei (Beechey ground squirrel).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
OC Spermophilus.
OX NCBI_TaxID=34862;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=95007566; PubMed=7923176;
RA Rivkina M.B., Cullen J.M., Robinson W.S., Marion P.L.;
RT "State of the p53 gene in hepatocellular carcinomas of ground
RT squirrels and woodchucks with past and ongoing infection with
RT hepadnaviruses.";
RL Cancer Res. 54:5430-5437(1994).
CC -1- FUNCTION: ACTS AS A TUMOR SUPPRESSOR IN MANY TUMOR TYPES. INDUCES
CC CIRCUMSTANCES OR APOPTOSIS DEPENDING ON THE PHYSIOLOGICAL
CC GROWTH ARREST OR CELL TYPE, BUT BOTH ACTIVITIES ARE INVOLVED IN
CC TUMOR SUPPRESSION. ACTS IN CELL CYCLE REGULATION, IS A TRANS-
CC ACTIVATOR THAT ACTS TO NEGATIVELY REGULATE CELL DIVISION BY
CC CONTROLLING A SET OF GENES REQUIRED FOR THIS PROCESS. ONE OF THE
CC ACTIVATED GENES IS AN INHIBITOR OF CYCLIN-DEPENDENT KINASES.
CC APOPTOSIS INDUCTION SEEMS TO BE MEDIATED EITHER BY STIMULATION OF
CC BAX AND FAS ANTIGEN EXPRESSION, OR BY REPRESSION OF BCL-2
CC EXPRESSION.
CC -1- SUBUNIT: BINDS DNA AS AN HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- DISEASE: P53 IS FOUND IN INCREASED AMOUNTS IN A WIDE VARIETY
CC OF TRANSFORMED CELLS. P53 IS FREQUENTLY MUTATED OR INACTIVATED
CC IN MANY TYPES OF CANCER.
CC -----
CC -1- SIMILARITY: BELONGS TO THE P53 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: U43902; AA85628.1; -.
CC HSP: P04637; 1YCS.
CC InterPro: IPR002117; -.
CC DR PIR: P500870; P53; 1.
CC Pfam: PF00348; P53; 1.
CC KMW Anti-oncogene; DNA-binding; Transcription regulation; Activator;
CC Nuclear protein; Phosphorylation; Apoptosis.
CC FT NON_TER 1 1
CC FT DOMAIN <1 24 TRANSCRIPTION ACTIVATION (ACIDIC).
CC FT DNA_BIND 80 270 BY SIMILARITY.
CC FT DOMAIN 303 >314 OLIGOMERIZATION.
CC FT DOMAIN 289 301 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

```

837	4	36.4	383	1	OPS4_DROVI	P17646	dtrosophila	910	4	36.4	414	1	LMPI1_CHICK	P05300	gallus gall
838	4	36.4	384	1	5X18_HUMAN	P35713	homo sapien	911	4	36.4	415	1	HEM1_NEIGO	O9zhd6	neisseria gall
839	4	36.4	384	1	Y039_MYCBE	P47285	mycoplasma	912	4	36.4	415	1	HEM1_NEIMA	O9zhd5	neisseria m
840	4	36.4	386	1	FADH_YEAST	P32771	saccharomyc	913	4	36.4	415	1	HEM1_NEIMB	P56994	neisseria m
841	4	36.4	386	1	GATM_PIG	P50441	sus scrofa	914	4	36.4	416	1	CTG1_HAEIN	P44548	haemophilus
842	4	36.4	387	1	ADHA_BACSU	O05239	bacillus su	915	4	36.4	416	1	IDH_ECOLI	P08200	escherichia
843	4	36.4	387	1	INVO_TARBA	P24771	tarsius dan	916	4	36.4	416	1	R23B_MOUSE	P54728	mus musc
844	4	36.4	387	1	SAT_DEIRA	P56864	delinococcus	917	4	36.4	417	1	FTSA_PSEAE	P47203	pseudomonas
845	4	36.4	388	1	ARGJ_CONGL	O59280	cornebacte	918	4	36.4	417	1	LYOX_HUMAN	P28300	homo sapien
846	4	36.4	388	1	CZFI_CANAL	P28875	candida alb	919	4	36.4	417	1	MAC1_YEAST	P16263	saccharomyc
847	4	36.4	388	1	VE2_HPV29	P50772	human papil	920	4	36.4	417	1	ODO2_BACSU	P16263	bacillus su
848	4	36.4	389	1	YB82_YEAST	O04951	saccharomyc	921	4	36.4	418	1	FTS2_TREPA	O83405	treponema p
849	4	36.4	390	1	ADHB_BACSU	O05240	bacillus su	922	4	36.4	419	1	KAS2_STRCN	O02579	streptomyces
850	4	36.4	390	1	RIR2_MOUSE	P11157	mus musc	923	4	36.4	419	1	LPB1_NEIMA	P57038	neisseria m
851	4	36.4	391	1	BCHP_RHOCA	P26172	rhodobacter	924	4	36.4	419	1	LPB1_NEIMB	O05014	neisseria m
852	4	36.4	391	1	NIRE_PSEST	O52521	pseudomonas	925	4	36.4	419	1	MANA_PSEFL	P194324	pseudomonas
853	4	36.4	391	1	VS06_ROTGA	P26015	rotavirus (	926	4	36.4	421	1	ACDM_HUMAN	P11330	homo sapien
854	4	36.4	392	1	CG2B_CHLYR	P51987	chlorenchytra	927	4	36.4	421	1	ACDM_PIG	P41367	sus scrofa
855	4	36.4	392	1	HRCA_CHLMU	P54306	chlamydia m	928	4	36.4	421	1	FXJ1_MOUSE	O61660	mus musc
856	4	36.4	393	1	GC1M_MOUSE	P01869	mus musc	929	4	36.4	421	1	FXJ1_RAT	O63247	rattus norv
857	4	36.4	393	1	TRUA_MOUSE	O9wu56	mus musc	930	4	36.4	421	1	MYB2_PHYPA	P80073	physcomit
858	4	36.4	394	1	VGID_HSV11	O69091	herpes simp	931	4	36.4	421	1	SYS_THERH	P34945	thermus aqu
859	4	36.4	394	1	VGID_HSV1A	P36318	herpes simp	932	4	36.4	422	1	CG2A_SPI50	P04962	spisula sol
860	4	36.4	394	1	VGID_HSV1P	P37083	herpes simp	933	4	36.4	422	1	SYH_HAEIN	P43823	haemophilus
861	4	36.4	395	1	YCHF_HSVB	P36342	simian herp	934	4	36.4	423	1	GATM_HUMAN	P50440	homo sapien
862	4	36.4	395	1	YCHF_CAEBL	P19197	caenorhabdi	935	4	36.4	423	1	GATM_RAT	P50442	rattus norv
863	4	36.4	395	1	C121_MYCTU	O59571	mycobacteri	936	4	36.4	423	1	IDH_BACSU	P39126	bacillus su
864	4	36.4	396	1	EPSC_BURSO	P52641	burkholderi	937	4	36.4	423	1	TYD_HUMAN	SHIP1_YEAST	homo sapien
865	4	36.4	396	1	YD61_MYCTU	O11031	mycobacteri	938	4	36.4	423	1	ASP_ANCCA	P32223	saccharomyc
866	4	36.4	397	1	SEPL_MOUSE	O62170	mus musc	939	4	36.4	424	1	ASP_ANCCA	O16937	anchylostom
867	4	36.4	398	1	EF1G_CAEBL	P54142	caenorhabdi	940	4	36.4	424	1	IVD_RAT	P12007	rattus norv
868	4	36.4	398	1	ODO2_AZOVI	P20708	azotobacter	941	4	36.4	424	1	KAS2_STRCO	P23156	streptomyces
869	4	36.4	398	1	TAP_DROME	O18667	dtrosophila	942	4	36.4	425	1	IDH_HELPJ	O9znd3	helicobacte
870	4	36.4	400	1	NICM_PROWIT	O37619	protobeca	943	4	36.4	425	1	IDH_HELPJ	P56603	helicobacte
871	4	36.4	401	1	NODE_RHIT	P04684	rhizodium 1	944	4	36.4	425	1	LMP2_CHICK	O90617	gallus gall
872	4	36.4	402	1	ARGJ_METUA	O57645	m putative	945	4	36.4	426	1	DHMH_PARVE	P23006	paracoccus
873	4	36.4	402	1	CAR1_CANPA	P32951	candida par	946	4	36.4	426	1	GSA_ECOLI	P23893	escherichia
874	4	36.4	402	1	HEM2_CHICK	O42479	gallus gall	947	4	36.4	426	1	GSA_SALTY	P42836	saccharomyc
875	4	36.4	402	1	KPPR_SPIOL	P09559	spinaecia ol	948	4	36.4	426	1	HEM1_XANCH	P42808	xanthomella
876	4	36.4	402	1	OPS4_CANAL	P46596	candida alb	949	4	36.4	427	1	CYB6_CHILT	O59297	chlorobium
877	4	36.4	402	1	PAI1_MOUSE	P22777	mus musc	950	4	36.4	427	1	TI31_HUMAN	P78552	homo sapien
878	4	36.4	402	1	PAI1_RAT	P20961	rattus norv	951	4	36.4	427	1	YBHC_ECOLI	P46130	escherichia
879	4	36.4	402	1	PGL1_ERMCA	P18192	erwinia car	952	4	36.4	428	1	HEM2_HORVU	O42836	hordeum vul
880	4	36.4	402	1	RDS1_SCHPO	P53693	schizosacch	953	4	36.4	428	1	MOD5_YEAST	P04884	saccharomyc
881	4	36.4	403	1	CGB3_CHICK	P33963	gallus gall	954	4	36.4	428	1	SCK_HUMAN	P98077	homo sapien
882	4	36.4	403	1	G3PA_MAIZE	P09315	zea mays (m	955	4	36.4	429	1	EP1G_ARPSA	P12261	artemia sal
883	4	36.4	403	1	I23O_HUMAN	P14902	homo sapien	956	4	36.4	429	1	MYCS_RAT	P23999	rattus norv
884	4	36.4	405	1	CAR8_CANAL	O42778	candida alb	957	4	36.4	429	1	YVFO_BACSU	O07013	bacillus su
885	4	36.4	405	1	GDPB_MOUSE	O921w4	mus musc	958	4	36.4	430	1	CINA_MYCTU	O07731	mycobacteri
886	4	36.4	405	1	THRC_METUA	O58860	methanococ	959	4	36.4	430	1	CREI_NEUCR	O59598	neurospora
887	4	36.4	405	1	YIS1_STRCO	P19780	streptomyces	960	4	36.4	430	1	GNL1_HUMAN	P36915	homo sapien
888	4	36.4	405	1	YUJ3_CAEBL	P34397	caenorhabdi	961	4	36.4	430	1	GSA_BACSU	P30949	bacillus su
889	4	36.4	407	1	COAT_BBV	P04329	black beeti	962	4	36.4	430	1	GSA_HELPJ	O9zmd0	helicobacte
890	4	36.4	407	1	COAT_FHV	P12870	flock house	963	4	36.4	430	1	GSA_HELPJ	P56115	helicobacte
891	4	36.4	407	1	CPXD_AGRU	P24467	agrobacteri	964	4	36.4	431	1	OAPA_HAEIN	P44415	haemophilus
892	4	36.4	408	1	GDPB_HUMAN	O59390	homo sapien	965	4	36.4	431	1	P2X6_HUMAN	O15547	homo sapien
893	4	36.4	408	1	YJ95_YEAST	P47160	saccharomyc	966	4	36.4	432	1	CGA2_HUMAN	P20248	homo sapien
894	4	36.4	409	1	HEM1_PARDE	P43089	paracoccus	967	4	36.4	432	1	ENV2_MOUSE	P11370	mus musc
895	4	36.4	409	1	KAS2_STRGA	P15539	streptomyces	968	4	36.4	432	1	TNU_RAT	P19332	rattus norv
896	4	36.4	409	1	R23B_HUMAN	P54727	homo sapien	969	4	36.4	433	1	TIG_HAEIN	P44837	haemophilus
897	4	36.4	410	1	CPXI_BACME	P14762	bacillus me	970	4	36.4	433	1	GAG_HTLV2	P03343	human t-cell
898	4	36.4	410	1	ENV_FLYCA	O02076	feline leuk	971	4	36.4	433	1	HEM2_SPIOL	P24493	spinaecia ol
899	4	36.4	410	1	FI82_ADEG1	O64762	avian adeno	972	4	36.4	433	1	HMR2_HAIVA	P42258	halobacteri
900	4	36.4	411	1	OS9C_DROME	O9u1p7	dtrosophila	973	4	36.4	433	1	RTCI_DICDI	O15746	dictyostell
901	4	36.4	412	1	T3RE_BACCE	P23241	bacillus ce	974	4	36.4	433	1	TNBI_AGRU	P54917	agrobacteri
902	4	36.4	412	1	Y360_MYCPN	P75241	mycoplasma	975	4	36.4	434	1	INXE_CAEBL	O62136	caenorhabdi
903	4	36.4	413	1	ESCI_SCHPO	O04635	schizosacch	976	4	36.4	434	1	PRSA_YEAST	P33297	saccharomyc
904	4	36.4	413	1	HMFT_DROME	P02835	dtrosophila	977	4	36.4	434	1	V61O_METUA	O58027	methanococ
905	4	36.4	413	1	OR7A_DROME	O9u315	dtrosophila	978	4	36.4	436	1	ENV_FLYCS	O02077	feline leuk
906	4	36.4	413	1	OST4_CHICK	P46440	gallus gall	979	4	36.4	436	1	RNO_AOUAE	O07683	halobacteri
907	4	36.4	413	1	PKK_DROME	P91622	dtrosophila	980	4	36.4	436	1	STD_HAUNI	V010_HSV6U	human herpe
908	4	36.4	413	1	YSO1_YEAST	P45819	saccharomyc	981	4	36.4	436	1			
909	4	36.4	414	1	IDH1_VIBAL	P41560	vibrio sp.	982	4	36.4	437	1	AAC3_DICDI	P14197	dictyostell

691	4	36.4	337	1	HAIR_DROME	P14003	drosophila	764	4	36.4	363	1	TOBI_MOUSE	O61471	mus musculus
692	4	36.4	337	1	VGH_BPC4	P03647	bacterioph	765	4	36.4	364	1	LIP_BURCE	P22088	burkholderi
693	4	36.4	337	1	YDTR_SCHPO	O14216	saccharosch	766	4	36.4	364	1	LIP_PSR55	P22275	pseudomonas
694	4	36.4	338	1	NH52_CAEEL	O17928	caenorhabdl	767	4	36.4	364	1	MRP_AOURB	O66946	aquifex aeo
695	4	36.4	338	1	SED1_YRST	O01389	saccharomyc	768	4	36.4	364	1	TAU2_MOUSE	P10638	mus musculus
696	4	36.4	338	1	TAP4_HUMAN	O01664	homo sapien	769	4	36.4	364	1	ZRP4_MAIZE	P47917	zea mays (m
697	4	36.4	339	1	E281_ARCFU	O29877	archaeoglob	770	4	36.4	365	1	MEFB_HUMAN	O02080	homo sapien
698	4	36.4	339	1	FMRE_DROVI	P18766	drosophila	771	4	36.4	365	1	OMLA_ACTPL	O02937	actinobacill
699	4	36.4	339	1	SRPA_SYNP7	O55025	synchococc	772	4	36.4	365	1	TRUA_HUMAN	O97606	homo sapien
700	4	36.4	340	1	CAC4_HUMAN	O11201	h cmp-n ace	773	4	36.4	366	1	CYCR_RHOGE	P51758	rhodocyclus
701	4	36.4	340	1	CATR_CHICK	P43223	gallus-gall	774	4	36.4	366	1	HNLS_SORBI	P52708	sorghum bic
702	4	36.4	340	1	RECA_MYCGE	P47581	mycoplasma	775	4	36.4	366	1	VF72_HAEIN	P46465	haemophilus
703	4	36.4	340	1	VP35_EBOZM	O05127	ebola virus	776	4	36.4	367	1	DNAT_COXBU	P42381	coxsiella bu
704	4	36.4	341	1	COA2_BFDV	P13892	budgerigar	777	4	36.4	368	1	GNUN_ECCLI	P37651	eschერიცხა
705	4	36.4	341	1	TAU3_MOUSE	P10637	mus musculu	778	4	36.4	368	1	MLTA_VIBCH	O9K994	vibriol choi
706	4	36.4	342	1	TONB_PSEAE	O51368	pseudomonas	779	4	36.4	368	1	MYCL_MOUSE	P10166	mus musculu
707	4	36.4	342	1	VU84_HSV6U	P52532	human herpe	780	4	36.4	368	1	SH31_HUMAN	O99961	homo sapien
708	4	36.4	343	1	COAL_BFDV	P13891	budgerigar	781	4	36.4	368	1	VE2_HPV6A	O84294	human papil
709	4	36.4	343	1	TRA_CORDI	P35879	corynebacte	782	4	36.4	368	1	YBHR_ECCLI	P44087	haemophilus
710	4	36.4	344	1	ASTA_ECCLI	P76218	eschერიცხა	783	4	36.4	368	1	Y974_HAEIN	P75774	eschერიცხა
711	4	36.4	344	1	CHEB_THEMA	O9WYN9	thermotoga	784	4	36.4	369	1	LMXB_MESAU	O60564	mesocricetu
712	4	36.4	344	1	HRCA_BACST	O45550	bacillus st	785	4	36.4	370	1	HMC2_DESVH	P33389	desulfocibtr
713	4	36.4	344	1	NUN2_RHIME	P56911	thizobium m	786	4	36.4	370	1	HUPK_RHILV	F28183	thizobium l
714	4	36.4	344	1	Y401_HUMAN	O43151	homo sapien	787	4	36.4	370	1	TRPD_MYCTU	Q10382	mycobacteri
715	4	36.4	344	1	YAAA_SCHPO	O09801	schizosacch	788	4	36.4	370	1	V2R_PIG	P32307	sus scrofa
716	4	36.4	345	1	GDPE_RAT	O92217	rattus norv	789	4	36.4	372	1	ADHA_RANPE	O57380	rana perezi
717	4	36.4	345	1	NU2M_MOUSE	P03893	mus musculu	790	4	36.4	372	1	COAL_POVMK	P24595	mouse polyo
718	4	36.4	345	1	NU2M_RAT	P11662	rattus norv	791	4	36.4	372	1	DBPA_HUMAN	P16989	homo sapien
719	4	36.4	345	1	PANC_YEAST	P40439	saccharomyc	792	4	36.4	372	1	LMXB_MOUSE	O88609	mus musculu
720	4	36.4	346	1	ARCC_BACSU	P23715	bacillus su	793	4	36.4	372	1	MSSP_HUMAN	P29558	homo sapien
721	4	36.4	346	1	BIOB_ECCLI	P12996	eschერიცხა	794	4	36.4	373	1	MP91_POAPR	P22284	poa pratens
722	4	36.4	346	1	BIOB_SERMA	P36569	serratia ma	795	4	36.4	374	1	GDPE_BRARE	O42222	brachydanio
723	4	36.4	346	1	T2D7_YEAST	O04254	saccharomyc	796	4	36.4	375	1	ACRG_EMENTI	P20309	emeritella
724	4	36.4	346	1	YIC3_YEAST	P40564	saccharomyc	797	4	36.4	375	1	ADHU_GADMO	P81600	gadus morhu
725	4	36.4	347	1	DDL_HELPJ	O92485	helicobacte	798	4	36.4	375	1	FENS_TOBAC	O04337	nicotiana t
726	4	36.4	347	1	DDL_HELPY	P56191	helicobacte	799	4	36.4	375	1	GDPE_BOVIN	O18836	bos taurus
727	4	36.4	347	1	DHAS_LEGPN	O31219	legionella	800	4	36.4	375	1	GDPE_CHICK	O42220	gallus gall
728	4	36.4	347	1	UL88_HSV7J	P52364	human herpe	801	4	36.4	375	1	GDPE_HUMAN	O14221	melagris g
729	4	36.4	349	1	SAPD_HAEIN	P45288	haemophilus	802	4	36.4	375	1	GDPE_METGA	O42223	melagris g
730	4	36.4	350	1	FLAO_TREPA	P18193	treponema p	803	4	36.4	375	1	GDPE_PAPHA	O18828	papio hamad
731	4	36.4	350	1	FLAO_HALYO	P41198	halobacteri	804	4	36.4	375	1	GDPE_PIG	O18831	sus scrofa
732	4	36.4	352	1	NU2M_STRPU	P15349	strongyloce	805	4	36.4	375	1	GDPE_SHEEP	O18830	ovis aries
733	4	36.4	352	1	SMP_SERNA	O06517	serratia ma	806	4	36.4	375	1	INTE_ECCLI	P75969	eschერიცხა
734	4	36.4	353	1	HEM2_BRAJA	P45622	bradyrhizob	807	4	36.4	375	1	OTC_CORHI	P78605	coriolus hi
735	4	36.4	353	1	YF91_MYCPN	O50336	mycoplasma	808	4	36.4	375	1	SRE_CHICK	O90717	gallus gall
736	4	36.4	354	1	YMA3_MYCBO	O02279	mycobacteri	809	4	36.4	376	1	DMPEM_STRIP	P42712	gallus gall
737	4	36.4	355	1	NIPS_LACDE	P31672	lactobacilli	810	4	36.4	376	1	VASP_MOUSE	P70460	mus musculu
738	4	36.4	355	1	VALI_ABRVW	P21947	abutillon mo	811	4	36.4	376	1	VE2_HPV10	P36781	human papil
739	4	36.4	355	1	Y644_SYNY3	O55724	synchocyst	812	4	36.4	377	1	FENS_PEA	O41014	pisum sativ
740	4	36.4	358	1	KLFI_MOUSE	P46099	mus musculu	813	4	36.4	377	1	LMXB_CHICK	P53413	gallus gall
741	4	36.4	358	1	TRMU_SYNY3	P73755	synchocyst	814	4	36.4	378	1	DHA_BACSU	O08352	bacillus su
742	4	36.4	358	1	VALI_BCVT	P14991	beet curly	815	4	36.4	378	1	FENS_ORYSA	P41345	oryza sativ
743	4	36.4	359	1	HIS8_SALTY	P10369	salmonella	816	4	36.4	378	1	FENS_ORYSA	O23877	oryza sativ
744	4	36.4	359	1	TMOD_HUMAN	P28289	homo sapien	817	4	36.4	378	1	MRG_HUMAN	P35410	homo sapien
745	4	36.4	359	1	TMOD_MOUSE	P49813	mus musculu	818	4	36.4	378	1	OPS4_DROME	O08255	drosophila
746	4	36.4	359	1	VALI_TYLCM	P27260	tomato yell	819	4	36.4	378	1	SVY_HAEPA	P36432	haemophilus
747	4	36.4	359	1	VALI_TYLCU	P38609	tomato yell	820	4	36.4	378	1	VE2_HPV30	P36790	human papil
748	4	36.4	360	1	GPRP_HUMAN	O00155	homo sapien	821	4	36.4	378	1	VE2_HPV55	O80937	human papil
749	4	36.4	360	1	VG47_BRPMU	O91132	bacterioph	822	4	36.4	379	1	LMXB_HUMAN	O60663	homo sapien
750	4	36.4	361	1	DBPA_RAT	O62764	rattus norv	823	4	36.4	379	1	P2X6_RAT	P51579	rattus norv
751	4	36.4	361	1	DCAM_HELAN	O65354	heliathus	824	4	36.4	379	1	VAT1_TORCA	P19333	torpedo cal
752	4	36.4	361	1	OSHI_ORYSA	P46609	oryza sativ	825	4	36.4	379	1	CAGD_YEAST	P28004	saccharomyc
753	4	36.4	361	1	RTCI_DROME	O77264	drosophila	826	4	36.4	380	1	CAGD_MOUSE	O64609	mus musculu
754	4	36.4	362	1	UFED1_YEAST	P53044	saccharomyc	827	4	36.4	380	1	VGLI_HSVB5	O08102	bovine heip
755	4	36.4	362	1	CKRA_MOUSE	P46092	homo sapien	828	4	36.4	380	1	YMH_CAEEL	P34474	caenorhabdl
756	4	36.4	362	1	HUPK_MOUSE	O9J121	mus musculu	829	4	36.4	381	1	GRK1_ECCLI	P77364	eschერიცხა
757	4	36.4	362	1	HUPK_BRAJA	P48342	bradyrhizob	830	4	36.4	382	1	LYSI_CANAL	P43065	candida alb
758	4	36.4	362	1	RLPA_ECCLI	P10100	eschერიცხა	831	4	36.4	382	1	MSNI_YEAST	P22148	saccharomyc
759	4	36.4	363	1	VM25_BORHE	P32779	borrella he	832	4	36.4	382	1	NU2M_CHLRE	P08740	chlamydomon
760	4	36.4	363	1	ADA_HUMAN	P00813	homo sapien	833	4	36.4	382	1	OP53_DROPS	P28660	drosophila
761	4	36.4	363	1	AGLK_RHIME	O92379	rhizobium m	834	4	36.4	383	1	RA27_YEAST	P26793	saccharomyc
762	4	36.4	363	1	R23A_MOUSE	P54725	homo sapien	835	4	36.4	383	1	HOXY_AUCEU	P31914	alcaligenes
763	4	36.4	363	1	R23A_MOUSE	P54726	mus musculu	836	4	36.4	383	1	METR_MYCPN	P78003	mycoplasma



545	4	36.4	272	1	THIM_BACSU	P39593	618	4	36.4	309	1	CBX3_RHOSH	P95648	rhodobacter
546	4	36.4	273	1	MYBC_MAIZE	P10290	619	4	36.4	309	1	YESP_BACSU	P35159	rhodobacter
547	4	36.4	275	1	APAH_ACTAC	O52655	620	4	36.4	310	1	YB84_HSV7J	O52534	human herpes
548	4	36.4	275	1	MADE_ECOLI	P18843	621	4	36.4	310	1	YB99_SCHPO	O13771	schizosacch
549	4	36.4	275	1	RK2_PTICAB	O62954	622	4	36.4	311	1	MENA_BACSU	P39582	rhodobacter
550	4	36.4	275	1	TF2D_ARTSF	O17488	623	4	36.4	311	1	MIAA_RICPR	O94d37	ricetella
551	4	36.4	277	1	T2FB_DROME	P41900	624	4	36.4	312	1	SNAG_HUMAN	O99747	homo sapien
552	4	36.4	277	1	VNST_CVBO	P18517	625	4	36.4	312	1	Y4K_RHISN	P55529	rhizobium s
553	4	36.4	278	1	HM37_CAEEL	O93356	626	4	36.4	314	1	GCR_SHEEP	P35547	ovis aries
554	4	36.4	278	1	HUPJ_RHCCA	O03009	627	4	36.4	314	1	MAG2_HUMAN	P43356	homo sapien
555	4	36.4	278	1	TRT3_RABIT	P02641	628	4	36.4	314	1	SSAK_SALTY	P78853	salmonella
556	4	36.4	279	1	UDP_KLEPN	P52671	629	4	36.4	315	1	CEP1_ENTHI	O01957	entamoeba h
557	4	36.4	280	1	MSA2_PLAFK	O03643	630	4	36.4	315	1	T2S1_STRAL	O53608	streptomyces
558	4	36.4	281	1	Y125_MYCPN	P75511	631	4	36.4	316	1	GRUP_BIOEL	O93134	blomphalar
559	4	36.4	282	1	GD43_WHEAT	P04723	632	4	36.4	316	1	LDH1_PLARD	O27743	plasmidum
560	4	36.4	283	1	Y150_ECOLI	P19769	633	4	36.4	316	1	YK95_MYCTU	O10704	mycobacteri
561	4	36.4	283	1	Y150_ECOLI	P32677	634	4	36.4	317	1	3M6_RAT	P23571	rattus norv
562	4	36.4	284	1	DRN1_MOUSE	P49183	635	4	36.4	317	1	AR34_SCHPO	O14241	schizosacch
563	4	36.4	284	1	DRN1_RAT	P21704	636	4	36.4	317	1	CBXC_ALCEU	P40118	alcaligenes
564	4	36.4	284	1	YPV2_METTF	P29571	637	4	36.4	317	1	CBXP_ALCEU	O04540	alcaligenes
565	4	36.4	285	1	PARP_ECOLI	P19071	638	4	36.4	317	1	CFXQ_XANFL	P23013	xanthobacte
566	4	36.4	285	1	T13B_HUMAN	O94275	639	4	36.4	317	1	FLGL_ECOLI	P28744	eschlerichia
567	4	36.4	286	1	CFXQ_CYACA	O22034	640	4	36.4	319	1	FTSN_ECOLI	P29131	eschlerichia
568	4	36.4	286	1	PROG_YEAST	P32263	641	4	36.4	319	1	YDFC_SCHPO	P04584	schizosacch
569	4	36.4	287	1	H1_LYCES	P37218	642	4	36.4	319	1	YDJP_BACSU	P54553	rhodobacter
570	4	36.4	288	1	PCD1_HUMAN	O15116	643	4	36.4	320	1	LDH_LACPE	P56511	lactobacill
571	4	36.4	288	1	PHNS_DESGI	P12943	644	4	36.4	320	1	YDIE_BACSU	P56512	lactobacill
572	4	36.4	288	1	SUCD_ECOLI	P07459	645	4	36.4	320	1	FLIM_AGRU	O34768	rhodobacter
573	4	36.4	288	1	VP30_EBOZM	O05323	646	4	36.4	321	1	ISP_BACCS	O44457	agrobacteri
574	4	36.4	288	1	Y134_MYCTU	O05099	647	4	36.4	321	1	ISP_BACCS	P29140	rhodobacter
575	4	36.4	289	1	BSN1_BACSU	O03091	648	4	36.4	321	1	PE14_ARATH	P48484	arabidopsis
576	4	36.4	290	1	ARY2_CHICK	P13914	649	4	36.4	321	1	YDGE_ECOLI	P21437	eschlerichia
577	4	36.4	290	1	Y444_METJA	O57886	650	4	36.4	322	1	LIPA_RHIRT	O05941	rhizobium e
578	4	36.4	291	1	CFXQ_ODOSI	P48826	651	4	36.4	322	1	PTNA_ECOLI	P08186	eschlerichia
579	4	36.4	291	1	EX53_MYGE	O49406	652	4	36.4	322	1	PLAU_LUPUL	P50345	lupinus lut
580	4	36.4	291	1	EX53_MYCPN	P73403	653	4	36.4	323	1	ANKH_CHRYT	O06527	chromatium
581	4	36.4	292	1	CGD3_MOUSE	P30282	654	4	36.4	323	1	GCL_RABIT	P01870	oryctolagus
582	4	36.4	293	1	CFXQ_GUTH	O78450	655	4	36.4	324	1	GCL_MOUSE	P01868	mus musculu
583	4	36.4	293	1	CGD3_RAT	P48961	656	4	36.4	324	1	PRGA_BACST	P42015	rhodobacter
584	4	36.4	293	1	SUCD_HAETN	P45102	657	4	36.4	326	1	GSFK_KLEPN	P17350	klebsiella
585	4	36.4	293	1	YF75_MYCTU	O50646	658	4	36.4	328	1	REBQ_NEIGO	P37363	neisseria g
586	4	36.4	294	1	LRR4_SYNPT	P52691	659	4	36.4	328	1	RPOA_BORPE	P37368	botetella
587	4	36.4	295	1	CFXQ_CYAME	O22025	660	4	36.4	328	1	RPOA_NEIMA	O91906	neisseria m
588	4	36.4	295	1	DHSB_USTMA	P32420	661	4	36.4	328	1	SNAG_BOVIN	P81127	homo taurus
589	4	36.4	295	1	MOTA_ECOLI	P09348	662	4	36.4	328	1	YOS1_CAEEL	O09461	caenorhabdi
590	4	36.4	295	1	MOTA_SALTY	P55891	663	4	36.4	329	1	CLP3_HUMAN	O15417	homo sapien
591	4	36.4	295	1	YFLE_CAUCR	P33976	664	4	36.4	329	1	RIR2_BACSU	P50621	rhodobacter
592	4	36.4	296	1	REBP2_YEAST	P03872	665	4	36.4	329	1	RL3_HA1HA	O06844	halobacteri
593	4	36.4	296	1	SAPR_STRPU	P11761	666	4	36.4	330	1	CLP3_RAT	P73797	rattus norv
594	4	36.4	296	1	TRA2_THIFE	P35884	667	4	36.4	330	1	EXL2_HUMAN	O04596	homo sapien
595	4	36.4	297	1	PECM_ERWCH	P42194	668	4	36.4	330	1	G3PC_LETME	O01558	leishmania
596	4	36.4	297	1	TF2D_XENIA	P27633	669	4	36.4	330	1	Y4KL_RHISN	P55530	rhizobium s
597	4	36.4	298	1	V084_HSV6	P27633	670	4	36.4	330	1	YHA9_YEAST	P38749	saccharomyc
598	4	36.4	299	1	H1S1_ECOLI	P52533	671	4	36.4	331	1	APAZ2_KLUDA	P49348	kluyveromyc
599	4	36.4	299	1	ICE1_SPOFR	P10366	672	4	36.4	331	1	LH15_EUGER	P08975	euglena gra
600	4	36.4	300	1	VMAT_MEASY	P36356	673	4	36.4	331	1	SUCA_SCHPO	O13750	schizosacch
601	4	36.4	301	1	CFXQ_PORPU	P51228	674	4	36.4	331	1	Y013_CAEEL	P34669	caenorhabdi
602	4	36.4	301	1	MCAT_HUMAN	O43772	675	4	36.4	332	1	COX2_SYNR3	O06474	synecocyst
603	4	36.4	301	1	MCAT_RAT	P97521	676	4	36.4	332	1	DNH1_SYNR3	P50027	synecocyst
604	4	36.4	301	1	SC14_KLUDA	P24859	677	4	36.4	332	1	GBLP_DICDI	P46800	dictyostell
605	4	36.4	301	1	YCSA_BACSU	P42958	678	4	36.4	332	1	HITA_HAETN	P33755	haemophilus
606	4	36.4	302	1	CLIB_LEUMC	O53078	679	4	36.4	333	1	3M6_MOUSE	P22885	mus musculu
607	4	36.4	302	1	COX2_RHOSH	O03736	680	4	36.4	333	1	MP92_POAPR	P22885	poa pratens
608	4	36.4	302	1	UI49_VZVD	P09272	681	4	36.4	333	1	OTC2_ECOLI	P06960	eschlerichia
609	4	36.4	303	1	VG08_BRP22	P26748	682	4	36.4	333	1	PLSX_BACSU	P71018	rhodobacter
610	4	36.4	304	1	CIK3_DROME	P08512	683	4	36.4	335	1	EBA2_FLAME	P336912	flavobacter
611	4	36.4	305	1	ATNB_CHICK	P08251	684	4	36.4	335	1	G3PC_CHOOR	P34920	chondrus cr
612	4	36.4	305	1	OAF_DROVI	O18638	685	4	36.4	335	1	MPFL_KLUDA	P34920	kluyveromyc
613	4	36.4	306	1	ISPE_MYCTU	O05956	686	4	36.4	335	1	OTCC_PSEAE	P08308	pseudomonas
614	4	36.4	306	1	SPBB_ECOLI	P10936	687	4	36.4	336	1	RLAU_MEYVA	P15826	methanococc
615	4	36.4	307	1	ALYA_KLEPN	O59478	688	4	36.4	337	1	YNC9_CAEEL	P34542	caenorhabdi
616	4	36.4	307	1	MP93_POAPR	P22286	689	4	36.4	337	1	CAG4_MOUSE	P54751	m cmp-n-ace
617	4	36.4	308	1	VPL_HAETN	P44227	690	4	36.4	337	1	G3P_COLLN	P54117	colletoetric



399	4	36.4	209	1	RL3_BACHD	092914 bacillus ha	472	4	36.4	242	1	MTGA_KLEPN	048465 klebsiella
400	4	36.4	209	1	RL3_BACSU	P42920 bacillus su	473	4	36.4	242	1	PDXJ_AQUAE	067171 aquifex ae
401	4	36.4	209	1	YC02_KLEPN	P48448 klebsiella	474	4	36.4	242	1	YD05_SCHPO	013725 schizosacch
402	4	36.4	210	1	NIP3_CAEEL	009969 caenorhabdi	475	4	36.4	244	1	SCS2_YEAST	P40075 saccharomyc
403	4	36.4	210	1	YACE_HAEIN	P44520 haemophilus	476	4	36.4	245	1	HI_MAIZE	P23444 zea mays (m
404	4	36.4	210	1	YCFO_ECOLI	P75920 escherichia	477	4	36.4	245	1	YM29_MYCTU	010513 mycobacteri
405	4	36.4	211	1	MAUD_METME	050332 methylophil	478	4	36.4	246	1	HEM4_CHLVI	059335 chlorobium
406	4	36.4	211	1	TUB8_SOLTU	P33191 solanum tub	479	4	36.4	246	1	MTGA_HAEIN	P44690 haemophilum
407	4	36.4	211	1	YFDX_ECOLI	P76520 escherichia	480	4	36.4	246	1	MYF5_COTJA	P34061 escherichia
408	4	36.4	211	1	YFYS_SCHPO	P36586 schizosacch	481	4	36.4	246	1	YBRC_ECOLI	P24371 escherichia
409	4	36.4	212	1	RERC_ARATH	092417 arabidopsis	482	4	36.4	246	1	YBRC_HAEIN	P44634 haemophilus
410	4	36.4	212	1	RL14_HUMAN	P50914 homo sapien	483	4	36.4	247	1	ATPI_CHLUV	P56295 chlorella v
411	4	36.4	212	1	SODP_VITVI	065199 vitis vinif	484	4	36.4	247	1	C21U_HUMAN	09umt2 homo sapien
412	4	36.4	213	1	HL3_RABIT	P02251 oryctolagus	485	4	36.4	247	1	CFI_RAPSA	022551 rephanus sa
413	4	36.4	213	1	RL14_RAT	063507 rattus norv	486	4	36.4	247	1	GRAB_HUMAN	P10144 h garyzme
414	4	36.4	213	1	RL3_BACST	P28600 bacillus st	487	4	36.4	247	1	YECO_ECOLI	P76290 escherichia
415	4	36.4	213	1	YCFM_ECOLI	P75947 escherichia	488	4	36.4	248	1	AMPW_MYCPN	011132 mycoplasma
416	4	36.4	214	1	YBFT_YEAST	P34222 saccharomyc	489	4	36.4	248	1	SOD1_PLEBO	001352 drosophila
417	4	36.4	214	1	YPU5_RHOCA	P26161 rhodobacter	490	4	36.4	249	1	BTUD_ECOLI	P50058 plectonema
418	4	36.4	215	1	RL25_MYCTU	P96385 mycobacteri	491	4	36.4	249	1	YXEO_BACSU	P06611 escherichia
419	4	36.4	215	1	RS6_MYCPN	P75543 mycoplasma	492	4	36.4	250	1	HXB9_MOUSE	P54854 bacillus su
420	4	36.4	216	1	HIC1_XENLA	P15866 xenopus lae	493	4	36.4	250	1	YXEO_MOUSE	P20615 mus musculu
421	4	36.4	217	1	OTCI_SALTY	008016 salmonella	494	4	36.4	251	1	RA1L_ARATH	Q38942 arabidopsis
422	4	36.4	217	1	YB47_YEAST	P38306 saccharomyc	495	4	36.4	252	1	END4_MAYLE	P30770 mycobacteri
423	4	36.4	217	1	YD67_SCHPO	010319 schizosacch	496	4	36.4	252	1	MYBD_MAIZE	P23592 zea mays (m
424	4	36.4	218	1	SCOB_MYCTU	006166 mycobacteri	497	4	36.4	252	1	TRT3_COTJA	P06539 coturnix co
425	4	36.4	219	1	ATPD_CHLRE	042687 chlamydomon	498	4	36.4	253	1	ADPL_DROME	P05552 drosophila
426	4	36.4	219	1	BASP_RAT	005175 rattus norv	499	4	36.4	254	1	PMFD_PROMI	P53520 proteus mlt
427	4	36.4	219	1	HLB_XENLA	P06693 xenopus lae	500	4	36.4	254	1	PT22_YEAST	P10355 saccharomyc
428	4	36.4	219	1	P1MT_PYRAB	09uxx0 pyrococcus	501	4	36.4	255	1	RS3A_CANAL	P40910 candida alb
429	4	36.4	220	1	HL3_MOUSE	P43277 mus musculu	502	4	36.4	255	1	SOLA_YEAST	P53315 saccharomyc
430	4	36.4	220	1	HIC2_XENLA	P15867 xenopus lae	503	4	36.4	255	1	TONB_YEREN	005740 yersinia en
431	4	36.4	221	1	HIC1_CHITE	P40277 chironomus	504	4	36.4	255	1	Y166_METJA	057630 methanococc
432	4	36.4	222	1	CASB_CAPHI	P33048 capra hircu	505	4	36.4	256	1	CB22_SOTBN	P09155 glycine max
433	4	36.4	222	1	CASB_SHEEP	P11839 ovis aries	506	4	36.4	256	1	FIXA_ECOLI	P31373 escherichia
434	4	36.4	223	1	AOP2_RAT	035244 r antioxiida	507	4	36.4	256	1	TATC_HAEIN	P44560 haemophilus
435	4	36.4	223	1	VTAT_LAMB	P03730 bacterioph	508	4	36.4	257	1	BF41_MOUSE	P28565 mus musculu
436	4	36.4	224	1	CASB_BOVIN	P02666 bos taurus	509	4	36.4	257	1	MODA_ECOLI	P33729 escherichia
437	4	36.4	224	1	RL4_CHLPN	092768 chlamydia p	510	4	36.4	257	1	TRT3_HUMAN	P45372 homo sapien
438	4	36.4	226	1	BASP_BOVIN	P80723 bos taurus	511	4	36.4	258	1	MYP5_CHICK	008856 gallus gall
439	4	36.4	226	1	BASP_HUMAN	P19752 neurospora	512	4	36.4	258	1	TRT3_RAT	P09379 rattus norv
440	4	36.4	228	1	HS30_NEUCR	09ztia4 rhodobacter	513	4	36.4	258	1	UL49_HSVBC	P30022 bovine herp
441	4	36.4	228	1	LEXA_RHOSE	080324 pyrus pyrif	514	4	36.4	260	1	1433_PEA	P46366 plasmu saliv
442	4	36.4	229	1	RNSE_PYRPP	027949 bison bonas	515	4	36.4	260	1	MTKB_METEX	P53595 methyllobact
443	4	36.4	229	1	SRY_BISBO	003255 bos taurus	516	4	36.4	260	1	TNFS_CANFA	097625 canis famli
444	4	36.4	229	1	SRY_BOVIN	004662 streptococc	517	4	36.4	260	1	TNFS_FELCA	097605 felis silve
445	4	36.4	230	1	CPSE_STRAG	P21895 chironomus	518	4	36.4	260	1	UPES_BACSU	031751 bacillus su
446	4	36.4	231	1	HL_CHITH	003700 bothriechis	519	4	36.4	260	1	VP33_APLCA	Q16943 aplysia cal
447	4	36.4	231	1	NU4M_BOTLA	09tvd0 camelus dro	520	4	36.4	261	1	YK07_YEAST	P36639 saccharomyc
448	4	36.4	232	1	CASB_CAMDR	P39037 sus scrofa	521	4	36.4	261	1	RS4_DROME	P41042 drosophila
449	4	36.4	232	1	CASH_PIG	055309 synchococc	522	4	36.4	261	1	YK83_CAEEL	P34349 caenorhabdi
450	4	36.4	232	1	CYSH_SYNP7	P40264 glyptotendi	523	4	36.4	262	1	Y182_CAEEL	P34440 caenorhabdi
451	4	36.4	232	1	HL2_GLYBA	P40275 chironomus	524	4	36.4	262	1	CORT_SHVX	004572 shalloe vit
452	4	36.4	232	1	HLA_CHITE	P40276 glyptotendi	525	4	36.4	262	1	FDHD_BACSU	P39156 bacillus su
453	4	36.4	232	1	H1B_CHITE	P00760 campylobact	526	4	36.4	264	1	TRT3_CHICK	P12620 gallus gall
454	4	36.4	232	1	TONB_CAMCO	P40263 glyptotendi	527	4	36.4	264	1	MSA2_PLAUF	P19599 plasmodium
455	4	36.4	233	1	H11_GLYBA	P40266 glyptotendi	528	4	36.4	264	1	MSA2_PLAUF	P50499 plasmodium
456	4	36.4	233	1	H11_GLYSA	092112 mus musculu	529	4	36.4	264	1	OTP_DROME	P56672 drosophila
457	4	36.4	233	1	SN64_MOUSE	092112 mus musculu	530	4	36.4	266	1	REF2_RAT	063528 rattus norv
458	4	36.4	234	1	ARGR_HUMAN	P55145 homo sapien	531	4	36.4	267	1	MM21_YEAST	P38633 saccharomyc
459	4	36.4	234	1	SN64_HUMAN	095473 homo sapien	532	4	36.4	267	1	RS3_MYCGE	P47403 mycoplasma
460	4	36.4	234	1	YER0_YEAST	P40011 saccharomyc	533	4	36.4	268	1	EKBA_AERHY	008437 aeromonas h
461	4	36.4	234	1	YGIB_ECOLI	P24195 escherichia	534	4	36.4	269	1	EXBD_SYNY3	P72942 synechocyst
462	4	36.4	235	1	Y101_YEAST	P40449 saccharomyc	535	4	36.4	269	1	PYR5_PREDI	P11401 firemyella d
463	4	36.4	236	1	PDX1_AUOPY	P46212 aquifex pyr	536	4	36.4	270	1	FKBA_ECOLI	P45523 escherichia
464	4	36.4	237	1	DCOP_LACPL	P77888 lactobacill	537	4	36.4	270	1	ISPD_STRCO	Q910q8 streptomyce
465	4	36.4	237	1	HET1_ANASP	P37695 anabaena sp	538	4	36.4	270	1	PANB_HELPJ	092598 helicobacte
466	4	36.4	238	1	RK5_ODOSI	P49547 odontella s	539	4	36.4	270	1	PEAZ_HELPY	025698 helicobacte
467	4	36.4	238	1	Y939_HAEIN	P44080 haemophilus	540	4	36.4	270	1	REF2_HUMAN	P15927 homo sapien
468	4	36.4	240	1	H12_VOICA	008865 volvox cart	541	4	36.4	270	1	REF2_MOUSE	062193 mus musculu
469	4	36.4	240	1	SFSA_METH	027565 methanobact	542	4	36.4	270	1	YD11_SCHPO	010247 schizosacch
470	4	36.4	241	1	SPIR_SPIME	P21655 spiroplasma	543	4	36.4	271	1	DJ1A_ECOLI	P31680 escherichia
471	4	36.4	242	1	MTGA_ECOLI	P46022 escherichia	544	4	36.4	271	1	VA04_VARV	P33832 variola vir

253	4	36.4	133	1	RCSF_ECOLI	P28633	escheric	326	4	36.4	172	1	RL10_BRUAB	P41107	bruce
254	4	36.4	134	1	ANFB_HUMAN	P16860	homo sap	327	4	36.4	173	1	Y013_BPH1	P51715	bacteri
255	4	36.4	137	1	H2B4_MAIZE	P42120	zea mays (m	328	4	36.4	175	1	OE22_BRANA	P29111	brassica
256	4	36.4	138	1	GRIM_DROME	O24570	droso	329	4	36.4	176	1	AROK_MYCTU	P55014	mycobact
257	4	36.4	139	1	THI2_ECOLI	P3636	escheric	330	4	36.4	176	1	ET2_RAT	P29943	rat
258	4	36.4	140	1	18K2_MYCT	P46732	mycobact	331	4	36.4	177	1	IMP2_YEAS	P26972	sacchar
259	4	36.4	140	1	HYPI_PIST1	P52748	pisol	332	4	36.4	178	1	ET2_HUMAN	P20800	homo sap
260	4	36.4	140	1	YOIB_ECOLI	P36652	escheric	333	4	36.4	179	1	RK6_GUTH	P46908	quillarl
261	4	36.4	141	1	NEUV_RANES	P11858	rana escul	334	4	36.4	184	1	GRPE_CHLPN	O92849	chlamydia
262	4	36.4	141	1	YO10_MYCTU	P15180	mycobact	335	4	36.4	184	1	LSPA_MYCPN	P75488	mycoplasma
263	4	36.4	141	1	ZRAP_ECOLI	P32682	escheric	336	4	36.4	184	1	NOSG_THERH	P35872	thermus aqu
264	4	36.4	142	1	H2AX_HUMAN	P16104	homo sap	337	4	36.4	184	1	YKY7_YEAS	O02205	sacchar
265	4	36.4	142	1	H2AX_MOUSE	P27661	mus muscu	338	4	36.4	185	1	KAD_SYNY3	P73302	synecocyst
266	4	36.4	142	1	HBAA_ANGAN	P80945	anguilla an	339	4	36.4	185	1	RM05_MARPO	P26786	marichanta
267	4	36.4	143	1	CMG4_BACSU	P25956	bacillus su	340	4	36.4	186	1	RAD_SYNB6	P42793	arabidops
268	4	36.4	143	1	LG81_VICRA	P02232	vicia faba	341	4	36.4	186	1	MIE1_MUGCA	P22159	mu
269	4	36.4	143	1	VLI_BPV3	P50805	bovine pap	342	4	36.4	186	1	RL18_YEAS	P07279	sacchar
270	4	36.4	143	1	VLI1_ADE41	P23688	human adeno	343	4	36.4	187	1	RL18_ARATH	P42791	arabidops
271	4	36.4	144	1	CORA_MOUSE	O62266	mus muscu	344	4	36.4	188	1	RM05_MARPO	P68660	marichanta
272	4	36.4	145	1	H2B_CAPAN	O49118	capsicum an	345	4	36.4	189	1	RS7A_YEAS	P26786	marichanta
273	4	36.4	145	1	SODM_STRPA	O33756	streptococ	346	4	36.4	190	1	YBAY_ECOLI	P77717	sacchar
274	4	36.4	146	1	FUR_HAEIN	P44561	haemophilus	347	4	36.4	190	1	YCH3_SCHPO	P78855	schizosacch
275	4	36.4	146	1	YIAC_ECOLI	P37664	escheric	348	4	36.4	191	1	EX01_RHIME	O32928	rhizobium m
276	4	36.4	147	1	CYPA_CAEEL	P52017	caenorhabd1	349	4	36.4	191	1	PGHD_SHEEP	O9xsm0	ovis aries
277	4	36.4	147	1	H2B_GOSHI	O25582	gossypium h	350	4	36.4	191	1	SDC2_XENLA	P49414	xenopus lae
278	4	36.4	147	1	RL13_MYCLE	P38014	mycobact	351	4	36.4	191	1	YGL2_STRCO	P40182	streptomyce
279	4	36.4	147	1	RL13_MYCTU	O06260	mycobact	352	4	36.4	192	1	UBC1_CAEEL	P52478	caenorhabd1
280	4	36.4	150	1	H2B2_MAIZE	P30756	zea mays (m	353	4	36.4	192	1	VG49_BRMD2	O64239	mycobact
281	4	36.4	150	1	NNA4_BPT3	P20313	ze	354	4	36.4	193	1	H10_HUMAN	P07305	homo sap
282	4	36.4	150	1	NNA4_BPT7	P00806	bacteri	355	4	36.4	193	1	VANI_VIBAN	P74945	vibrio angu
283	4	36.4	150	1	TAAI_VACC	P20982	vac	356	4	36.4	193	1	Y457_METJA	O57897	methanococ
284	4	36.4	150	1	TAI1_VACCV	P07610	vac	357	4	36.4	194	1	HL_SALT	P02254	salmo trutt
285	4	36.4	150	1	TA02_VARY	P33814	var	358	4	36.4	194	1	MIEV_HUMAN	P08590	homo sap
286	4	36.4	150	1	TA02_SCHPO	O10081	schizosacch	359	4	36.4	195	1	COAG_LIMPO	P03998	limulus pol
287	4	36.4	151	1	H2B1_MAIZE	P30755	zea mays (m	360	4	36.4	196	1	ANOD_ARCFU	O30011	archaeoglob
288	4	36.4	151	1	RBSD_ECOLI	P04982	escheric	361	4	36.4	196	1	MOBA_STACA	O9zlm1	staphylococ
289	4	36.4	151	1	SYRB_RHIME	O33685	rhizobium m	362	4	36.4	196	1	RGS1_HUMAN	O08116	homo sap
290	4	36.4	152	1	ADOM_BPT3	P07693	bacteri	363	4	36.4	196	1	TRAP_ECOLI	P41068	escheric
291	4	36.4	152	1	CORA_RAT	O65532	rattus norv	364	4	36.4	197	1	SDC4_CHICK	P49416	gallus gall
292	4	36.4	152	1	H2B1_MHEAT	P27807	triticum ae	365	4	36.4	197	1	MRBA_ECOLI	P30849	escheric
293	4	36.4	153	1	CORB_MOUSE	O62267	mus muscu	366	4	36.4	197	1	YJ10_MYCTU	O07723	mycobact
294	4	36.4	153	1	H2B3_MAIZE	O43261	zea mays (m	367	4	36.4	198	1	CSP_BOVIN	O29455	bos taurus
295	4	36.4	153	1	PYRI_SERMA	P19936	serat	368	4	36.4	198	1	CSP_MOUSE	P4101	mus muscu
296	4	36.4	154	1	H2B5_MAIZE	P54348	zea mays (m	369	4	36.4	198	1	NUDC_EMENT	P17624	emer
297	4	36.4	155	1	YFH3_YEAS	P43587	sacchar	370	4	36.4	198	1	SDC4_MOUSE	P31431	homo sap
298	4	36.4	156	1	RL2B_HUMAN	P29316	homo sap	371	4	36.4	198	1	SDC4_MOUSE	O35988	mus muscu
299	4	36.4	157	1	RAP_TAROF	O49065	taraxacum o	372	4	36.4	198	1	YH13_YEAS	P38896	sacchar
300	4	36.4	157	1	RL24_HUMAN	P38663	homo sap	373	4	36.4	198	1	YH8G_PSEPU	P20162	pseudomonas
301	4	36.4	157	1	YM75_YEAS	O05027	sacchar	374	4	36.4	199	1	D522_CRAPL	O01931	craterostig
302	4	36.4	158	1	ILVH_BUCAI	P57320	buchnera ap	375	4	36.4	199	1	MLEV_RAT	P16409	rattus norv
303	4	36.4	158	1	ILVH_BUCAP	O85294	buchnera ap	376	4	36.4	199	1	PRL_TOXAF	P10765	toxodonta a
304	4	36.4	159	1	DRR1_PEA	P12329	pisum sativ	377	4	36.4	199	1	PSBQ_CHLRE	P12852	chlamydomon
305	4	36.4	160	1	ET2_MOUSE	P22389	mus muscu	378	4	36.4	201	1	GDIS_HUMAN	P25266	homo sap
306	4	36.4	160	1	RL10_METH	O27191	methanobact	379	4	36.4	201	1	IF3_MYCTU	P34975	mycobact
307	4	36.4	160	1	VIPC_MYCHR	P29230	mycoplasma	380	4	36.4	201	1	SDC2_HUMAN	P34741	homo sap
308	4	36.4	161	1	CAMP_BRAFL	O01305	branchiost	381	4	36.4	201	1	SDC2_RAT	P34900	rattus norv
309	4	36.4	161	1	CAVP_BRAVA	P04573	branchiost	382	4	36.4	201	1	YKDO_YEAS	P36099	sacchar
310	4	36.4	161	1	DH3_HORVU	P12948	hordeum vul	383	4	36.4	201	1	YPT1_PHTYN	O01890	phyto
311	4	36.4	161	1	HS13_ARATH	P19037	arabidops	384	4	36.4	202	1	SDC2_MOUSE	P43407	mus muscu
312	4	36.4	162	1	GVP1_HALME	O02234	halobacteri	385	4	36.4	202	1	SDC4_RAT	P34901	rattus norv
313	4	36.4	162	1	YB09_SCHPO	P87150	schizosacch	386	4	36.4	202	1	YIC2_AGRRH	P13460	agrobact
314	4	36.4	163	1	GSPM_AERHY	P41850	aeromonas h	387	4	36.4	202	1	YGE7_YEAS	P33178	sacchar
315	4	36.4	163	1	NTEB_HERSE	P27714	herbaspirill	388	4	36.4	203	1	BCRC_BACLI	P42334	bacillus li
316	4	36.4	164	1	PTPA_STRCO	P53433	streptomyc	389	4	36.4	203	1	HYCB_ECOLI	P16428	escheric
317	4	36.4	164	1	Y400_SYNY3	O35129	synecocyst	390	4	36.4	204	1	VATD_BORBU	O51119	bortella bu
318	4	36.4	167	1	PLMP_GRIFR	P81054	griffo	391	4	36.4	204	1	YMA44_CAEEL	P34449	caenorhabd1
319	4	36.4	167	1	TCPT_YEAS	P35691	sacchar	392	4	36.4	206	1	AMIS_RHOER	O33185	rhodococ
320	4	36.4	167	1	YIG5_YEAS	P40366	sacchar	393	4	36.4	206	1	HL_ONCMY	P06355	oncorhynch
321	4	36.4	168	1	GLP_MOUSE	P14220	mus muscu	394	4	36.4	206	1	KAD_AQUAE	O66499	aquifex ae
322	4	36.4	169	1	Z160_HUMAN	O14589	homo sap	395	4	36.4	207	1	COX3_BACSU	P24012	bacillus su
323	4	36.4	171	1	IF3_BACST	P03000	bacillus st	396	4	36.4	207	1	H1T_MACMO	P40286	macaca mula
324	4	36.4	171	1	Y008_CAEEL	P34646	caenorhabd1	397	4	36.4	208	1	FTS0_STRGR	P5503	streptomyc
325	4	36.4	172	1	ILVH_METJA	O57625	methanococ	398	4	36.4	208	1	V23K_PEBV	P16261	pea early b

107	5	45.5	714	1	FRDA_HELIPY	006913	helicobacte	180	4	36.4	105	1	APLI_PETMA	P07095	petromyzon
108	5	45.5	766	1	DAB2_MOUSE	P98078	mus musculu	181	4	36.4	105	1	LAC_HUMAN	P01442	homo sapien
109	5	45.5	770	1	DAB2_HUMAN	P98082	mus musculu	182	4	36.4	105	1	THIO_RHOSH	P08058	rhodobacter
110	5	45.5	788	1	IF41_WHEAT	003387	triticeum ae	183	4	36.4	105	1	VPR_HV2D1	P17761	human immun
111	5	45.5	800	1	INLA_LISMO	P25166	listeria mo	184	4	36.4	105	1	VPR_HV2G1	P18046	human immun
112	5	45.5	863	1	IMB1_SCPPO	013864	schizosacch	185	4	36.4	105	1	VPR_HV2N2	P05930	human immun
113	5	45.5	893	1	YM92_CAEEL	P34531	caenorhabdi	186	4	36.4	105	1	VPR_HV2RO	P06938	human immun
114	5	45.5	894	1	KDPO_ECOLI	P21865	escherichia	187	4	36.4	105	1	Y027_MYCTU	P71597	mycobacteri
115	5	45.5	1027	1	ISWI_DROME	024368	drosoophila	188	4	36.4	105	1	Y1F4_YEAST	P40524	saccharomyc
116	5	45.5	1040	1	MAN1_RAT	P21139	rattus norv	189	4	36.4	107	1	COL_RABIT	P42890	oryctolagus
117	5	45.5	1051	1	ULK1_MOUSE	070405	mus musculu	190	4	36.4	108	1	KVIR_HUMAN	P01610	homo sapien
118	5	45.5	1061	1	RNE_ECOLI	P21513	escherichia	191	4	36.4	108	1	THIO_ECOLI	P00274	escherichia
119	5	45.5	1070	1	AGLO_CANTS	P29054	candida tsu	192	4	36.4	109	1	LV10_HUMAN	P06888	homo sapien
120	5	45.5	1218	1	MGPC_MYCPN	050341	mycoplasma	193	4	36.4	109	1	NUOM_BOVIN	P25712	homo sapien
121	5	45.5	1301	1	PTP9_DROME	P35832	drosoophila	194	4	36.4	109	1	VP26_BPAPS	091522	bacteriophe
122	5	45.5	1403	1	PRO_DROME	P29617	drosoophila	195	4	36.4	111	1	IM98_CAEEL	091042	caenorhabdi
123	5	45.5	1552	1	TP2A_CHICK	042130	gallus gall	196	4	36.4	111	1	LVIC_HUMAN	P01702	homo sapien
124	5	45.5	1612	1	TP2B_CRILLO	064399	criceulius	197	4	36.4	111	1	LV1D_HUMAN	P01702	homo sapien
125	5	45.5	1612	1	TP2B_MOUSE	064511	mus musculu	198	4	36.4	111	1	LV2H_HUMAN	P01711	homo sapien
126	5	45.5	1663	1	HAPP_HUMAN	060229	homo sapien	199	4	36.4	112	1	COL_CANFA	P19090	cantis famli
127	5	45.5	1919	1	HAPP_RAT	P97924	rattus norv	200	4	36.4	112	1	LV1B_HUMAN	P01700	homo sapien
128	5	45.5	2134	1	Y192_HUMAN	093074	homo sapien	201	4	36.4	112	1	LV1H_HUMAN	P06887	homo sapien
129	5	45.5	2278	1	FAB1_YEAST	P34756	saccharomyc	202	4	36.4	113	1	KV2F_MOUSE	P01630	mus musculu
130	5	45.5	358	1	DYHC_CAEEL	019020	caenorhabdi	203	4	36.4	113	1	RSBV_STRCO	Q9VWX8	streptomyce
131	5	45.5	364	1	PYSB_METBA	P80522	methanosarc	204	4	36.4	113	1	SS12_STRRO	P29607	streptomyce
132	5	45.5	364	1	SCKG_PANIM	010726	pandinus im	205	4	36.4	114	1	RK20_CHUVU	P56352	chlorella v
133	5	45.5	364	1	NEUY_ONCMY	P29071	oncorhynch	206	4	36.4	114	1	RSBV_LISMO	085016	listeria mo
134	5	45.5	56	1	MGLEV_MOUSE	P09542	mus musculu	207	4	36.4	114	1	YHIT_BUCAI	P07683	buchnera ap
135	5	45.5	56	1	VG36_BPMD2	022001	mycobacteri	208	4	36.4	116	1	ANSR_BACSU	007468	bacillus su
136	5	45.5	67	1	VPX_BPP2	P51772	bacterioph	209	4	36.4	116	1	MCS_HUMAN	P49901	homo sapien
137	5	45.5	70	1	FMFL_AERHY	P19369	aeromonas h	210	4	36.4	116	1	RLJ3_CHLRE	P45841	chlamydomon
138	5	45.5	70	1	PSAE_NOSS8	09W91	nostoc sp.	211	4	36.4	117	1	RLJ3_CHLRE	P22952	chlamydomon
139	5	45.5	75	1	YEJL_ECOLI	P33921	escherichia	212	4	36.4	117	1	AMC2_PIG	046204	chlamydia p
140	5	45.5	79	1	ET2_MACRA	028470	macaca fasc	213	4	36.4	117	1	YMOI_CHLPS	P34485	caenorhabdi
141	5	45.5	81	1	HEMP_YEREN	P31516	yersinia en	214	4	36.4	118	1	ELI1_PHYCR	P41892	phytophthor
142	5	45.5	81	1	RLJ3_MERTH	027649	methanobact	215	4	36.4	118	1	ELI1_PHYCR	P15570	phytophthor
143	5	45.5	83	1	MULI_PSEAE	P11221	pseudomonas	216	4	36.4	118	1	ELI1_PHYCR	P41801	phytophthor
144	5	45.5	83	1	VPP_BPBF2	P27331	bacterioph	217	4	36.4	118	1	LHG2_RHOCA	P23430	rhodobacter
145	5	45.5	83	1	Y9KD_BPBF2	P19193	bacterioph	218	4	36.4	118	1	RL7_MICLU	P02395	micrococcus
146	5	45.5	85	1	COXG_BOVIN	P00429	bos taurus	219	4	36.4	119	1	B2MG_CHISA	007523	chitropotes
147	5	45.5	85	1	YRBA_HAETN	P45026	haemophilus	220	4	36.4	119	1	PA21_OXYS	P00614	oxytarrus s
148	5	45.5	87	1	TAT_CAEVG	P21125	caenophillus	221	4	36.4	119	1	YEAR_ECOLI	P76248	escherichia
149	5	45.5	87	1	VPR_HV2BE	P18100	human immun	222	4	36.4	120	1	RLJ3_NICGU	P46290	nicotiana g
150	5	45.5	87	1	YN55_CAEEL	P34589	caenorhabdi	223	4	36.4	120	1	RLJ3_PICMA	065071	picea maria
151	5	45.5	89	1	RS15_PSEPU	087151	pseudomonas	224	4	36.4	120	1	RL34_ARATH	042331	arabidopsis
152	5	45.5	90	1	HEO_HAETN	P44437	haemophilus	225	4	36.4	121	1	AMEL_ORNAN	097646	ornithothyrn
153	5	45.5	90	1	TEGP_HSVSB	P30025	simlan heip	226	4	36.4	121	1	H2B1_TERTH	P08993	tetrahymena
154	5	45.5	91	1	VE4_HPV58	P26349	human papil	227	4	36.4	121	1	H2B2_TERTH	P08994	tetrahymena
155	5	45.5	94	1	RUXE_YEAST	Q12330	saccharomyc	228	4	36.4	122	1	RL14_SYNY3	F73310	synecocyst
156	5	45.5	95	1	ADXH_DROME	P37193	dirosophila	229	4	36.4	123	1	ELI2_PHYCR	P41803	phytophthor
157	5	45.5	95	1	SM32_HUMAN	P55855	homo sapien	230	4	36.4	123	1	Y949_METJA	058359	methanococc
158	5	45.5	97	1	CSOA_THINE	P45689	thiobactilli	231	4	36.4	124	1	RL7_BORBU	051363	borrella bu
159	5	45.5	97	1	VG45_BPML5	005256	mycobacteri	232	4	36.4	124	1	RL7_BORBU	P41106	bruceella ab
160	5	45.5	98	1	COXB_BOVIN	P00428	bos taurus	233	4	36.4	125	1	PSAE_SPTOL	P12334	sphinctia ol
161	5	45.5	98	1	CSOC_THINE	P45688	thiobactilli	234	4	36.4	125	1	RP09_SCHPO	009171	schizosacch
162	5	45.5	98	1	ELIA_PHYCP	P15571	phytophthor	235	4	36.4	125	1	RS6_CAMJE	092913	campylobact
163	5	45.5	98	1	ELIA_PHYDR	P35686	phytophthor	236	4	36.4	125	1	Y364_AQUAE	066689	aquifex aeo
164	5	45.5	98	1	ELIA_PHYME	P35686	phytophthor	237	4	36.4	125	1	Y944_HELPY	P46964	helicobacte
165	5	45.5	98	1	ELIB_PHYCI	P15569	phytophthor	238	4	36.4	125	1	Y944_HELPY	025558	helicobacte
166	5	45.5	98	1	ELIB_PHYDR	P35687	phytophthor	239	4	36.4	126	1	US05_HCMVA	P16840	human cytom
167	5	45.5	98	1	ELIB_PHYME	P35689	phytophthor	240	4	36.4	127	1	RL20_STRCO	088003	streptomyce
168	5	45.5	100	1	V07K_NMY	P15098	narcissus m	241	4	36.4	128	1	H1SX_AZOBH	P18766	azospirilli
169	5	45.5	101	1	FLIO_ECOLI	P22508	escherichia	242	4	36.4	129	1	HLVX_HUMAN	P04432	homo sapien
170	5	45.5	101	1	LSM3_HUMAN	09V451	homo sapien	243	4	36.4	129	1	OSTE_YEAST	P46964	saccharomyc
171	5	45.5	101	1	VG25_BPMD2	064219	mycobacteri	244	4	36.4	130	1	LVIG_HUMAN	P06316	homo sapien
172	5	45.5	101	1	VG25_BPML5	Q05232	mycobacteri	245	4	36.4	130	1	YHLS_YEAST	P38793	saccharomyc
173	5	45.5	101	1	Y0XC_BACSU	P28670	bacillus su	246	4	36.4	131	1	DFRA_MYXXA	052178	myxococcus
174	5	45.5	102	1	NODM_RHLIT	052846	rhizobium i	247	4	36.4	131	1	FRDC_ECOLI	P03803	escherichia
175	5	45.5	103	1	LV1E_HUMAN	P01703	homo sapien	248	4	36.4	131	1	FRDC_PROVU	P20923	proteus vul
176	5	45.5	103	1	SM31_HUMAN	P55854	homo sapien	249	4	36.4	131	1	IL13_MOUSE	P20193	mus musculu
177	5	45.5	103	1	YGU3_YEAST	P53097	saccharomyc	250	4	36.4	132	1	TLA_CANFA	077762	cantis famli
178	5	45.5	104	1	VPR_HV2CA	P24111	human immun	251	4	36.4	132	1	TV42_MOUSE	R01729	mus musculu
179	5	45.5	104	1	VPR_HV2ST	P20884	human immun	252	4	36.4	132	1	Y0S5_ARCFU	030181	archaeoglob

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:42:54 ; Search time 20.38 seconds  
(without alignments)  
18.489 Million cell updates/sec

Title: US-09-372-036-29

Perfect score: 11

Sequence: 1 QOQTAPKAPTE 11

Scoring table: OLIGO

Searched: 93435 seqs, 34255486 residues

Word size: 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database: SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	11	100.0	PE0_LISMO	P21171 listeria mo
2	6	54.5	P53_SPEBE	Q64662 sperophilu
3	6	54.5	P53_MARMO	Q36006 marmota mon
4	6	54.5	IF2_SKYX3	P72689 synchocyst
5	6	54.5	NCC2_MOUSE	Q60591 mus musculu
6	6	54.5	RPRO_SMYEA	P28897 strabertry
7	6	54.5	ANT1_ONCYO	P21249 onchocerca
8	5	45.5	GSP1_AERYV	P31737 aeromonas h
9	5	45.5	CU19_ARADI	P80515 araneus dia
10	5	45.5	H2B_HOLTV	P48557 holothuria
11	5	45.5	WN3A_ALOYU	P28101 alopias vul
12	5	45.5	WN3A_MELGA	P28125 melagris g
13	5	45.5	WN3A_SCECC	P28142 sceloporius
14	5	45.5	CU24_ARADI	P80516 araneus dia
15	5	45.5	CU26_ARADI	P80517 araneus dia
16	5	45.5	ULB0_HCMVA	P16830 human cytom
17	5	45.5	MSC1_HAETN	P44789 haemophilus
18	5	45.5	CYC3_DESYH	P00131 desulfovibr
19	5	45.5	CYC3_DESYV	P00132 desulfovibr
20	5	45.5	Y142_LEUMC	P97117 leuconostoc
21	5	45.5	Y147_CAEEL	Q11074 caenorhabdi
22	5	45.5	DH1_HORVU	P12951 hordeum vul
23	5	45.5	GP27_BPSP1	P06228 bacteriopho
24	5	45.5	FLIN_PSEAB	Q51466 pseudomonas
25	5	45.5	H1_DICDI	P54671 dictyostell
26	5	45.5	YPOL_THETL	P32438 thermus aqu
27	5	45.5	NOO9_THERL	Q56224 thermus aqu
28	5	45.5	IF3_PSEFL	Q9x627 pseudomonas
29	5	45.5	IF3_PSESY	P52834 pseudomonas
30	5	45.5	H12_CAEEL	P157329 caenorhabdi
31	5	45.5	Y319_MYCPN	P17396 mycoplasma
32	5	45.5	KAPS_YEAST	Q02196 saccharomyc
33	5	45.5	KAD_YEREN	P43412 yersinia en

34	5	45.5	214	1	KAD_YERPE	Q69172 yersinia pe
35	5	45.5	223	1	TONB_XANCP	Q34261 xanthomonas
36	5	45.5	231	1	NAPG_ECOLI	P33936 escherichia
37	5	45.5	235	1	HIE_CHIPA	P40262 chironomus
38	5	45.5	236	1	H1_WHEAT	P27806 triticum ae
39	5	45.5	237	1	HIE_CHITE	P40278 chironomus
40	5	45.5	239	1	OPAA_NEIGO	Q04876 neisseria g
41	5	45.5	241	1	NOCM_AGRU	P33113 agrobacteri
42	5	45.5	244	1	H10_CHITH	Q07134 chironomus
43	5	45.5	252	1	NODE_MYCTU	P95177 mycobacteri
44	5	45.5	259	1	HXC9_FUGRU	Q42502 fugu rubrip
45	5	45.5	260	1	H11_VOLCA	Q08864 volvox cart
46	5	45.5	279	1	THET_THIEVU	P04072 thermactin
47	5	45.5	280	1	LXLI_HUMAN	P52954 homo sapien
48	5	45.5	288	1	LIPH_PSEAN	Q01725 pseudomonas
49	5	45.5	292	1	CDG3_HUMAN	P30281 homo sapien
50	5	45.5	296	1	SMDF_HUMAN	Q15491 homo sapien
51	5	45.5	298	1	XERC_ECOLI	P22885 escherichia
52	5	45.5	300	1	XERC_SALTY	P55888 salmonella
53	5	45.5	313	1	DRPE_CRAPL	P22242 craterostig
54	5	45.5	316	1	MANA_STRMU	O59935 streptococ
55	5	45.5	318	1	KITH_FLUV	P22176 fish lympho
56	5	45.5	319	1	SSUA_ECOLI	P75853 escherichia
57	5	45.5	320	1	TAUA_ECOLI	Q47537 escherichia
58	5	45.5	320	1	YD33_YEAST	Q12117 saccharomyc
59	5	45.5	321	1	Y377_MYCTU	O53712 mycobacteri
60	5	45.5	325	1	CYF_STNP2	P28293 synchococ
61	5	45.5	332	1	SR4_PHYPO	P11113 physarum po
62	5	45.5	340	1	LIMA_PSEAE	Q04591 pseudomonas
63	5	45.5	341	1	VP3_GELV	P17768 grapevine f
64	5	45.5	343	1	CAG4_PIG	Q02745 s cmp-n-ace
65	5	45.5	344	1	ILVC_LACLA	Q00138 lactococcus
66	5	45.5	362	1	YCHE_HAETN	P44681 haemophilus
67	5	45.5	380	1	VASP_HUMAN	P50552 homo sapien
68	5	45.5	398	1	DAP3_HUMAN	P51398 homo sapien
69	5	45.5	398	1	RA23_YEAST	P32628 saccharomyc
70	5	45.5	414	1	KAS2_STRHA	Q05357 streptomyce
71	5	45.5	415	1	KAS2_STRVN	P16541 streptomyce
72	5	45.5	448	1	PSA4_SCHPO	P36612 schizosacch
73	5	45.5	451	1	WORD_BACSU	Q03522 bacillus su
74	5	45.5	457	1	ODR7_CAEEL	P14933 caenorhabdi
75	5	45.5	470	1	DAX1_HUMAN	P51843 homo sapien
76	5	45.5	474	1	PX2_CAVPO	O70397 cavia porce
77	5	45.5	479	1	PK2_DICDI	P28178 dictyostell
78	5	45.5	481	1	PE0_LISIN	Q01836 listeria in
79	5	45.5	483	1	OPRM_PSEAE	O51487 pseudomonas
80	5	45.5	493	1	PDI_CHICK	P09102 gallus gall
81	5	45.5	508	1	GSBP_CHICK	P12244 gallus gall
82	5	45.5	508	1	PDI_HUMAN	P07237 homo sapien
83	5	45.5	509	1	PDI_MOUSE	P09103 mus musculu
84	5	45.5	509	1	PDI_RABIT	P21195 oryctolagus
85	5	45.5	509	1	PDI_RAT	P04785 rattus norv
86	5	45.5	510	1	BMP6_MOUSE	P207072 mus musculu
87	5	45.5	510	1	PDI_BOVIN	P05307 bos taurus
88	5	45.5	520	1	CP84_ARATH	Q42600 arabidopsi
89	5	45.5	523	1	PE0_LISSE	Q01838 listeria se
90	5	45.5	524	1	PE0_LISWE	Q01839 listeria we
91	5	45.5	537	1	CNE3_HUMAN	O75131 homo sapien
92	5	45.5	546	1	RUB1_BRANA	P12244 brassica na
93	5	45.5	550	1	CCE_DROME	P14046 drosophila
94	5	45.5	553	1	ODO2_MYCTU	O10381 mycobacteri
95	5	45.5	554	1	SYNC_YEAST	P38707 saccharomyc
96	5	45.5	557	1	MERA_SHEPU	O54465 shewanella
97	5	45.5	583	1	ACES_BOVIN	P23795 bos taurus
98	5	45.5	584	1	YMB3_YEAST	Q09128 saccharomyc
99	5	45.5	598	1	PEX5_YARLI	Q09144 yarrowia il
100	5	45.5	624	1	NIFA_AZOLI	P54929 azospirillum
101	5	45.5	637	1	DNAR_ECOLI	P04475 escherichia
102	5	45.5	638	1	YHES_HAETN	P44808 haemophilus
103	5	45.5	656	1	FRDA_WOLSU	P17412 wolfinella s
104	5	45.5	667	1	TS11_GIALA	O00185 giardia lam
105	5	45.5	681	1	VEL_HPV10	P33720 human papil
106	5	45.5	714	1	FRDA_HELPJ	Q92mpo helicobacte

**This Page Blank (uspio)**

## C:Function:

A:Description: catalyzes reversible phosphorylation of AMP with ATP to form two ADP  
A>Note: magnesium required

C:Superfamily: adenylate kinase

C:Keywords: ATP; P-loop; phosphotransferase

F:7-14/Region: nucleotide-binding motif A (P-loop) #status atypical

F:80-85/Region: nucleotide-binding motif B #status atypical

F:126/Active site: His #status predicted

## Query Match

45.5%; Score 5; DB 2; Length 214;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OOTAP 6

DB 173 OOTAP 177

## RESULT 50

T38907

hypothetical protein SPAC56E4.05 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 04-Mar-2000

C:Accession: T38907

R:Conor, R.; Church, C.M.; Barrell, B.G.; Raftandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, September 1997

A:Reference number: Z21813

A:Accession: T38907

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-216 <CON>

A:Cross-references: EMBL:Z99261; PIDN:CAR16396.1; GSPDB:GN00066; SPDB:SPAC56E4.05

A:Experimental source: strain 972h-; cosmid c56E4

C:Genetics:

A:Gene: SPDB:SPAC56E4.05

A:Map position: 1

C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPAC56E4.05

## Query Match

45.5%; Score 5; DB 2; Length 216;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 KAPTE 11

DB 196 KAPTE 200

Search completed: August 15, 2001, 12:34:40  
Job time: 105 sec

RESULT 45  
H81057  
C:Species: Neisseria meningitidis  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: H81057  
R:Retelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; rt, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizzi, M.  
Science 287, 1809-1815, 2000  
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappelli, R.; V  
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.  
A:Reference number: AB1000, PMID:2015755  
A:Accession: H81057  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-205 <TET>  
A:Cross-references: GB:AE002516; GB:AE002098; NID:g7226905; PIDN:AAF42006.1; PID:g722690  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB1657

Query Match 45.5%; Score 5; DB 2; Length 205;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
|||||  
Db 191 APRAP 195

RESULT 46  
T49218  
hypothetical protein F27H5.20 - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C:Accession: T49218  
R:Reger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K  
Submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z25018  
A:Accession: T49218  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-207 <RIE>  
A:Cross-references: EMBL:AL163852; GSPDB:GN00061; ATSP:F27H5.20  
C:Genetics:  
A:Experimental source: cultivar Columbia; BAC clone F27H5  
A:Gene: ATSP:F27H5.20  
A:Map position: 3  
A:Introns: 66/2; 104/1; 149/1; 169/3

Query Match 45.5%; Score 5; DB 2; Length 207;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PKAPT 10  
|||||  
Db 126 PKAPT 130

RESULT 47  
DB3944  
SOS regulon transcription repressor *lexA* [Imported] - *Bacillus halodurans* (strain C-125)  
C:Species: *Bacillus halodurans*  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 08-Dec-2000  
C:Accession: DB3944  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and

A:Reference number: AB3650; PMID:20263314  
A:Accession: DB3944  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-207 <STO>  
A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BAB06075.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: *lexA*

Query Match 45.5%; Score 5; DB 2; Length 207;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOOTA 5  
|||||  
Db 145 QOOTA 149

RESULT 48  
H75317  
translation initiation factor IF-3 - *Deinococcus radiodurans* (strain R1)  
C:Species: *Deinococcus radiodurans*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: H75317  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zaleski, C.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
A:Reference number: A75250; PMID:20036896  
A:Accession: H75317  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-212 <WHI>  
A:Cross-references: GB:AE002044; GB:AE000513; NID:g6459872; PIDN:AAF11633.1; PID:g645  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR2087  
A:Map position: 1  
C:Superfamily: translation initiation factor IF-3

Query Match 45.5%; Score 5; DB 2; Length 212;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
|||||  
Db 169 APRAP 173

RESULT 49  
S70734  
adenylate kinase (EC 2.7.4.3) - *Yersinia enterocolitica*  
N:Alternate names: ATP-AMP transphosphorylase  
C:Species: *Yersinia enterocolitica*  
C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 18-Jun-1999  
C:Accession: S70734; S51258  
R:Skurnik, M.; Venho, R.; Toivanen, P.; Al-Hendy, A.  
Mol. Microbiol. 17, 575-594, 1995  
A:Title: A novel locus of *Yersinia enterocolitica* serotype O:3 involved in lipopolysa  
A:Reference number: S70734; PMID:96100456  
A:Accession: S70734  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-214 <SKU>  
A:Cross-references: EMBL:Z47767; NID:g633689; PIDN:CAAB7696.1; PID:g633690  
A:Experimental source: strain 6471/76 serotype O:3  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995  
C:Genetics:  
A:Gene: *ack*

R.White, O.; Elsen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A'Reference number: AF5250; MUID:20036896  
 A'Accession: E75435  
 A>Status: preliminary  
 A'Molecule type: DNA  
 A'Residuals: 1-183 <WHI>  
 A'Cross-references: GB:AE001961; GB:AE000513; NID:g6458643; PIDN:AAF10683.1; PID:g645864  
 A'Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR1110  
 A:Map position: 1  
 C:Superfamily: hypoxanthine phosphoribosyltransferase

Query Match 45.5%; Score 5; DB 2; Length 183;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
 |||||  
 DB 2 TAPKA 6

RESULT 42  
 T42232  
 histone H1.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 04-Mar-2000  
 C'Accession: T42232; S08195  
 R:Jedrusik, M.; Schulze, E.  
 submitted to the EMBL Data Library, August 1997  
 A'Description: The histone H1 complement of Caenorhabditis elegans.  
 A'Reference number: 222091  
 A'Accession: T42232  
 A>Status: translated from GB/EMBL/DBJ  
 A'Molecule type: mRNA  
 A'Residuals: 1-191 <JED>  
 A'Cross-references: EMBL:AF017811; NID:g2407322; PIDN:AAB70666.1; PID:g2407323  
 R:Vanfleteren, J.R.; Van Bun, S.M.; de Baere, I.; van Beeumen, J.J.  
 Biochem. J. 265, 739-746, 1990  
 A>Title: The primary structure of a minor isoform (H1.2) of histone H1 from the nematode  
 A'Reference number: S08195; MUID:90165873  
 A'Accession: S08195  
 A'Molecule type: protein  
 A'Residuals: 2-27, 'T', '29-30, 'SS', '33-44, 'IKE', '48-49, 'KQ', '52-78, 'R', '80-111, 'P', '113-191 <VAN  
 A>Note: the sequence from Fig. 5 is inconsistent with that from Fig. 4 in having 95-Arg  
 C:Superfamily: histone H1  
 C:Keywords: blocked amino end; chromosomal protein; DNA binding; nucleosome; nucleus  
 F:2-191/Product: histone H1.2 #status predicted <MAT>  
 F:2/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #statu

Query Match 45.5%; Score 5; DB 2; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 DB 18 APKAP 22

RESULT 43  
 S73713  
 MG319 homolog H08-orf193 - Mycoplasma pneumoniae (strain ATCC 29342)  
 C:Species: Mycoplasma pneumoniae  
 A:Variety: ATCC 29342  
 C>Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
 C'Accession: S73713  
 R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkil, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996  
 A>Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumon  
 A'Reference number: S73327; MUID:97105885  
 A'Accession: S73713  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A'Molecule type: DNA  
 A'Residuals: 1-193 <HIM>  
 A'Cross-references: EMBL:AE000037; GB:U00089; NID:g1674065; PIDN:AAB96035.1; PID:g167  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
 C:Genetics:  
 A:Genetic code: SGC3

Query Match 45.5%; Score 5; DB 2; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
 |||||  
 DB 68 TAPKA 72

RESULT 44  
 S17244  
 adenylylsulfate kinase (EC 2.7.1.25) - Yeast (Saccharomyces cerevisiae)  
 N:Alternate names: adenosine-5-phosphosulfate kinase; protein YKL001c  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 12-Nov-1999  
 C'Accession: S17244; S25323; S37811  
 R:Korich, C.; Mountain, H.A.; Byström, A.S.  
 Mol. Gen. Genet. 229, 96-108, 1991  
 A>Title: Cloning, nucleotide sequence, and regulation of MET14, the gene encoding the  
 A'Reference number: S17244; MUID:91375456  
 A'Accession: S17244  
 A'Molecule type: DNA  
 A'Residuals: 1-202 <KOR>  
 A'Cross-references: GB:S55315; NID:g235814; PIDN:AAB19854.1; PID:g235815  
 R:Duesterhoeft, A.; Philippson, P.  
 Yeast 8, 749-759, 1992  
 A>Title: DNA sequencing and analysis of a 24.7 kb segment encompassing centromere CEN  
 A'Reference number: S25323; MUID:93070612  
 A'Accession: S25323  
 A>Status: nucleic acid sequence not shown  
 A'Molecule type: DNA  
 A'Residuals: 1-202 <DUE>  
 A'Cross-references: EMBL:X65124; NID:g3517; PIDN:CAA46252.1; PID:g3529  
 A'Experimental source: strain S288C  
 R:Duesterhoeft, A.; Moestl, D.; Poehlmann, R.; Philippson, P.  
 submitted to the Protein Sequence Database, March 1994  
 A'Reference number: S37811  
 A'Accession: S37811  
 A'Molecule type: DNA  
 A'Residuals: 1-202 <DUO>  
 A'Cross-references: EMBL:Z28001; NID:g485972; PIDN:CAA81833.1; PID:g485973; GSPDB:GNO  
 A'Experimental source: strain S288C  
 C:Genetics:  
 A:Gene: SGD:MET14; MIPS:YKL001C  
 A'Cross-references: SGD:S0001484; MIPS:YKL001C  
 A:Map position: 11L  
 C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology  
 C:Keywords: ATP; P-loop; phosphotransferase; purine nucleotide binding  
 F:24-187/Domain: adenylylsulfate kinase homology <ASK>  
 F:31-38/Region: nucleotide-binding motif A (P-loop)  
 F:37/Binding site: ATP (Lys) #status predicted

Query Match 45.5%; Score 5; DB 1; Length 202;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 DB 166 APKAP 170



A:Molecule type: DNA  
A:Residues: 1-174 <RAM>  
A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BAA31056.1; PID:g3258373  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Gene: PH1929

Query Match 45.5%; Score 5; DB 2; Length 174;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
|1111|  
DB 23 APRAP 27

RESULT 37  
A26882  
p12 hypothetical protein - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 05-Nov-1999  
C:Accession: A26882  
R:Glatchenhaus, N.; Cuzin, F.  
Cell 50, 1081-1089, 1987  
A:Title: A role for ID repetitive sequences in growth and transformation-dependent regul  
A:Reference number: A26882; MID:87301751  
A:Accession: A26882  
A:Molecule type: mRNA  
A:Residues: 1-175 <GLA>  
A:Cross-references: GB:M17412; NID:g207249; PIDN:AAA4223.1; PID:g207250  
A:Experimental source: fibroblast  
A:Note: the authors translated the codon TGG for residue 65 as Ser and GAA for residue 1

Query Match 45.5%; Score 5; DB 2; Length 175;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PKAPT 10  
|1111|  
DB 39 PKAPT 43

RESULT 38  
T48699  
hypothetical protein IAG.30 [imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 19-May-2000  
C:Accession: T48699  
R:Schulte, U.; Aign, V.; Hohlisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24541  
A:Accession: T48699  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-176 <SCH>  
A:Cross-references: EMBL:AL353817; GSPDB:GN00112; NCSP:IAG.30  
A:Experimental source: cosmid contig IAG; strain 74  
C:Genetics:  
A:Gene: NCSP:IAG.30  
A:Map position: 2  
C:Superfamily: Neurospora crassa hypothetical protein IAG.30

Query Match 45.5%; Score 5; DB 2; Length 176;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OQTAP 6  
|1111|  
DB 153 OQTAP 157

## RESULT 39

F70588  
hypothetical protein Kv2843 - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: F70588  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Radandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; MID:98255987  
A:Accession: F70588  
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-181 <COL>  
A:Cross-references: GB:Z95207; GB:AL123456; NID:g3261745; PIDN:CAB08447.1; PID:e31518  
C:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: Kv2843

Query Match 45.5%; Score 5; DB 2; Length 181;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
|1111|  
DB 41 APRAP 45

RESULT 40  
T11906  
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain N009 - Thermus aquaticus thermophil  
N:Alternate names: NADH dehydrogenase I, subunit N009  
C:Species: Thermus aquaticus thermophilus  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 17-Mar-2000  
C:Accession: T11906  
R:Yano, T.; Chu, S.S.; Sled, V.S.; Ohnishi, T.; Yagi, T.  
submitted to the EMBL Data Library, March 1996  
A:Description: Cloning, sequencing, and expression studies of the proton-translocati  
A:Reference number: Z17372  
A:Accession: T11906  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-182 <YAN>  
A:Cross-references: EMBL:U52917; NID:g1279860; PID:g1279869; PIDN:AAA57946.1  
A:Experimental source: strain HB-8  
C:Genetics:  
A:Gene: N009  
C:Superfamily: unassigned ferredoxin 2(4Fe-4S)-related proteins; ferredoxin 2(4Fe-4S)  
C:Keywords: NAD: oxidoreductase  
F:46-116/Domain: ferredoxin 2(4Fe-4S) homology <FER>

Query Match 45.5%; Score 5; DB 2; Length 182;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 KAPTE 11  
|1111|  
DB 174 KAPTE 178

RESULT 41  
E75435  
pyrimidine operon regulatory protein PyrR - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: E75435

A:Experimental source: cv. Georgie  
C:Superfamily: dehydrin-like protein

Query Match 45.5%; Score 5; DB 2; Length 139;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 00QTA 5  
|||||  
DB 102 00QTA 106

## RESULT 32

C81214

comEA-related protein NMB0299 [imported] - Neisseria meningitidis (strain MC58 serogroup C)  
Species: Neisseria meningitidis  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: C81214

R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; et al.; Qiu, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizza, M.  
Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; et al.  
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A:Reference number: AB1000; MUID:20157555

A:Accession: C81214

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-148 &lt;TER&gt;

A:Cross-references: GB:AE002387; GB:AE002098; NID:g7225523; PIDN:AAFA0750.1; PID:g722552  
A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0299

Query Match 45.5%; Score 5; DB 2; Length 148;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
|||||  
DB 134 APKAP 138

## RESULT 33

F83464

flagellar motor switch protein FljN PA1444 [imported] - Pseudomonas aeruginosa (strain F)  
Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83464

R:Stover, C.K.; Pham, X.Q.; Ertvin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim, J.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: AB2950; MUID:20437337

A:Accession: F83464

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-157 <STO>  
A:Cross-references: GB:AE004574; GB:AE004091; NID:g9947391; PIDN:AA04833.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: FljN; PA1444

Query Match 45.5%; Score 5; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PKAPT 10  
|||||

DB 62 PKAPT 66

## RESULT 34

F81813

hypothetical protein NMA1867 [imported] - Neisseria meningitidis (strain 22491 serogr  
Species: Neisseria meningitidis  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: F81813

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo  
; Holroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre  
Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491  
A:Reference number: AB1775; MUID:20222556

A:Accession: F81813

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-159 &lt;PAR&gt;

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CA85090.1; PID:g738  
A:Experimental source: serogroup A, strain 22491

C:Genetics:

A:Gene: NMA1867

Query Match 45.5%; Score 5; DB 2; Length 159;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QTPAK 7  
|||||  
DB 81 QTPAK 85

## RESULT 35

I39546

hypothetical protein 2 - Aeromonas hydrophila  
Species: Aeromonas hydrophila  
C:Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 08-Oct-1999  
C:Accession: I39546

R:Thomas, S.R.; Trust, T.J.  
J. Bacteriol. 177, 3932-3939, 1995

A:Title: A specific Pili homolog is required for the secretion of paracrystalline sur  
A:Reference number: A57354; MUID:95332195

A:Accession: I39546

A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A:Residues: 1-166 &lt;RES&gt;

A:Cross-references: GB:LA1682; NID:g950300; PIDN:AAA79321.1; PID:g1019922  
C:Superfamily: Aeromonas hydrophila hypothetical protein 2

Query Match 45.5%; Score 5; DB 2; Length 166;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
|||||  
DB 26 TAPKA 30

## RESULT 36

A71208

hypothetical protein PH1929 - Pyrococcus horikoshii  
Species: Pyrococcus horikoshii  
C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
C:Accession: A71208

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu  
DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A:Reference number: A71000; MUID:98344137

A:Accession: A71208

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Cross-references: DDBJ:D31702; NID:g496361; PIDN:BAA06511.1; PID:g496362  
R:Shinkai, W.; Hase, T.; Yagi, T.; Matsubara, H.  
J. Biochem. 87, 1747-1756, 1980  
A;Title: Amino acid sequence of cytochrome c-3 from *Desulfovibrio vulgaris*, Miyazaki.  
A;Reference number: A00125; M0ID:8029474  
A;Accession: A00125  
A;Molecule type: protein  
A;Residues: 24-64, 'N', 66-130 <SH2>  
R:Higuchi, Y.; Kusunoki, M.; Matsuura, Y.; Yasuoka, N.; Kakudo, M.  
J. Mol. Biol. 172, 109-139, 1984  
A;Title: Refined structure of cytochrome c-3 at 1.8 angstrom resolution.  
A;Reference number: A49705; M0ID:84114880  
A;Contents: annotation; X-ray crystallography, 1.8 angstroms  
R:Higuchi, Y.; Kusunoki, M.; Matsuura, Y.; Yasuoka, N.; Kakudo, M.  
submitted to the Brookhaven Protein Data Bank, November 1983  
A;Reference number: A50415; PDB:2CDV  
A;Contents: annotation; X-ray crystallography, 1.8 angstroms, residues 24-64, 'N', 66-130  
C;Function:  
A;Description: accepts electrons from cytochrome-c3 hydrogenase (EC 1.12.2.1) and transfer  
A;Pathway: sulfate respiration  
C;Superfamily: cytochrome c3; cytochrome c3 homology  
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; sulfate respiration  
F:43-129/Product: cytochrome c3 status experimental <MAT>  
F:43-129/Domain: cytochrome c3 homology <CC3>  
F:45-57/Binding site: heme iron (His) (axial ligands) #status experimental  
F:48-106/Binding site: heme iron (His) (axial ligands) #status experimental  
F:53-56/Binding site: heme (Cys) (covalent) #status experimental  
F:58-75/Binding site: heme iron (His) (axial ligands) #status experimental  
F:93-129/Binding site: heme (Cys) (covalent) #status experimental  
F:93-129/Binding site: heme iron (His) (axial ligands) #status experimental  
F:102-105/Binding site: heme (Cys) (covalent) #status experimental  
F:123-128/Binding site: heme (Cys) (covalent) #status experimental

Query Match 45.5%; Score 5; DB 1; Length 130;  
Best Local Similarity 100.0%; Pred. No. 78;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
|||||  
Db 24 APRAP 28

RESULT 28  
H82438  
large-conductance mechanosensitive channel VCA0612 [imported] - *Vibrio cholerae* (strain  
C;Species: *Vibrio cholerae*  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C;Accession: H82438  
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.  
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A;Reference number: A82035; M0ID:20406833  
A;Accession: H82438  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-136 <HEI>  
A;Cross-references: GB:AE004391; GB:AE003853; NID:g9658015; PIDN:AAF96513.1; GSPDB:GN001  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VCA0612  
A;Map position: 2  
C;Superfamily: yhcC protein

Query Match 45.5%; Score 5; DB 2; Length 136;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAP 9  
|||||

Db 109 APRAP 113

RESULT 29  
S35752  
dehydrin 7 - barley  
C;Species: *Hordeum vulgare* (barley)  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 20-Aug-1999  
C;Accession: S35752  
R:Robertson, M.; Close, T.J.; Cumling, A.C.  
submitted to the EMBL Data Library, April 1993  
A;Description: Sequence analysis of a dehydration-induced gene, dehydrin, from barley  
A;Reference number: S35752  
A;Accession: S35752  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-137 <ROB>  
A;Cross-references: EMBL:X71362; NID:g296197; PIDN:CAA50499.1; PID:g296198  
C;Genetics:  
A;Intons: 66/73  
C;Superfamily: dehydrin-like protein

Query Match 45.5%; Score 5; DB 2; Length 137;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOQRA 5  
|||||  
Db 100 QOQRA 104

RESULT 30  
S05544  
dehydrin 8 - barley  
N;Alternate names: dehydration-induced protein 8  
C;Species: *Hordeum vulgare* (barley)  
C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 20-Aug-1999  
C;Accession: S05544  
R:Close, T.J.; Kortt, A.A.; Chandler, P.M.  
Plant Mol. Biol. 13, 95-108, 1989  
A;Title: A cDNA-based comparison of dehydration-induced proteins (dehydrins) in barley  
A;Reference number: S05543; M0ID:93357436  
A;Accession: S05544  
A;Molecule type: mRNA  
A;Residues: 1-139 <CIO>  
A;Cross-references: EMBL:X15288; NID:g18965; PIDN:CAA33362.1; PID:g18966  
C;Superfamily: dehydrin-like protein

Query Match 45.5%; Score 5; DB 2; Length 139;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QOQRA 5  
|||||  
Db 102 QOQRA 106

RESULT 31  
T05715  
dehydrin - barley  
C;Species: *Hordeum vulgare* (barley)  
C;Date: 09-Apr-1999 #sequence\_revision 09-Apr-1999 #text\_change 20-Jun-2000  
C;Accession: T05715  
R:Gull, M.  
submitted to the EMBL Data Library, June 1996  
A;Reference number: Z15427  
A;Accession: T05715  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-139 <GUL>  
A;Cross-references: EMBL:X98326; PIDN:CAA66970.1

A:Molecule type: DNA  
A:Residues: 1-127 <CHES>  
A:Cross-references: EMBL:X17403; NID:959591; PIDN:CAA35348.1; PID:e27313; PID:g1780891  
A>Note: this sequence was submitted to the EMBL Data Library, December 1989  
C:Superfamily: human cytomegalovirus hypothetical protein UL110

Query Match 45.5%; Score 5; DB 2; Length 127;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OOTAP 6  
|||||  
DB 23 OOTAP 27

## RESULT 24

A72468  
hypothetical protein APE2544 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C>Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 08-Sep-2000  
C:Accession: A72468  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix  
A:Reference number: A72450; MUID:99310339  
A:Accession: A72468  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-127 <KAM>  
A:Cross-references: DDBJ:AP000064; NID:g5105945; PIDN:BAA81561.1; PID:d1045347; PID:g510  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE2544  
C:Superfamily: Aeropyrum pernix hypothetical protein APE2544

Query Match 45.5%; Score 5; DB 2; Length 127;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
|||||  
DB 80 TAPKA 84

## RESULT 25

G64155  
yhdc protein - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 29-Sep-1999  
C:Accession: G64155  
R:Flisbach, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shiley, R.; Liu, L.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: G64155  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-128 <TIGR>  
A:Cross-references: GB:U02745; GB:I42023; NID:g1573617; PIDN:AAC22286.1; PID:g1573622; C:Genetics:  
A:Gene: yhdc  
C:Superfamily: yhdc protein

Query Match 45.5%; Score 5; DB 2; Length 128;  
Best Local Similarity 100.0%; Pred. No. 77;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 4 TAPKA 8  
|||||  
DB 110 TAPKA 114

## RESULT 26

CCDV3  
cytochrome c3 precursor - Desulfovibrio vulgaris (strain Hildenborough)  
C:Species: Desulfovibrio vulgaris  
C>Date: 24-Apr-1984 #sequence\_revision 17-Feb-1994 #text\_change 03-Mar-2000  
C:Accession: A24799; A00124; D32427  
R:Voordouw, G.; Brenner, S. Eur. J. Biochem. 159, 347-351, 1986  
A:Title: Cloning and sequencing of the gene encoding cytochrome c3 from Desulfovibrio  
A:Reference number: A24799; MUID:87004646  
A:Accession: A24799  
A:Molecule type: DNA  
A:Residues: 1-129 <VOO>  
A:Cross-references: GB:X04304; NID:g40820; PIDN:CAA27847.1; PID:g40821  
R:Troun, E.B.; Campbell, L.L. J. Biol. Chem. 249, 386-393, 1974  
A:Title: Amino acid sequence of cytochrome c-3 from Desulfovibrio vulgaris.  
A:Reference number: A00124; MUID:74070664  
A:Accession: A00124  
A:Molecule type: protein  
A:Residues: 23-129 <TRO>  
A:Experimental source: strain Hildenborough, NCIB 8303  
R:Loutfi, M.; Guerlesquin, F.; Bianco, P.; Haladjian, J.; Brusch, M. Biochem. Biophys. Res. Commun. 159, 670-676, 1989  
A:Title: Comparative studies of polymeric cytochromes c isolated from Desulfovibrio v  
A:Reference number: A32427; MUID:89193654  
A:Accession: D32427

A:Molecule type: protein  
A:Residues: 23-44 <LOU>  
C:Superfamily: cytochrome c3; cytochrome c3 homology  
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; sulfate res  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-129/Product: cytochrome c3 #status experimental <KAM>  
F:44-128/Domain: cytochrome c3 homology <CC3>  
F:44,56/Binding site: heme iron (His) (axial ligands) #status predicted  
F:47,105/Binding site: heme iron (His) (axial ligands) #status predicted  
F:52,55/Binding site: heme (Cys) (covalent) #status predicted  
F:57,74/Binding site: heme iron (His) (axial ligands) #status predicted  
F:68,73/Binding site: heme (Cys) (covalent) #status predicted  
F:92,128/Binding site: heme iron (His) (axial ligands) #status predicted  
F:101,104/Binding site: heme (Cys) (covalent) #status predicted  
F:122,127/Binding site: heme (Cys) (covalent) #status predicted

Query Match 45.5%; Score 5; DB 1; Length 129;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
|||||  
DB 23 APKAP 27

## RESULT 27

CCDV3M  
cytochrome c3 precursor [validated] - Desulfovibrio vulgaris (strain Miyazaki)  
C:Species: Desulfovibrio vulgaris  
C>Date: 31-Oct-1980 #sequence\_revision 10-Oct-1997 #text\_change 15-Sep-2000  
C:Accession: S33874; A00125  
R:Kikimura, M.; Ozawa, K.; Kojima, S.; Kunagai, I.; Akutsu, H.; Miura, K. Protein Seq. Data Anal. 5, 193-196, 1993  
A:Title: The primary structure of pre-cytochrome c(3) from Desulfovibrio vulgaris (M  
A:Reference number: S33874  
A:Accession: S33874  
A:Molecule type: DNA  
A:Residues: 1-130 <KIT>

A;Cross-references: EMBL:X71782; NID:g417605; PIDN:CAA50667.1; PID:g417606

Query Match 45.5%; Score 5; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APKAP 9  
|||||  
DB 63 APKAP 67

## RESULT 19

Hit family protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: G75374  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: G75374

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-118 <WHI>

A:Cross-references: GB:AE002005; GB:AE000513; NID:g6459377; PIDN:AAF11181.1; PID:g645938

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1621

A:Map position: 1

C:Superfamily: protein kinase C inhibitor; histidine triad homology

Query Match 45.5%; Score 5; DB 2; Length 118;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APKAP 9  
|||||  
DB 40 APKAP 44

## RESULT 20

B49905

protein secretion operon exe protein I precursor - Aeromonas hydrophila

C:Species: Aeromonas hydrophila

C>Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 02-Sep-2000

C:Accession: B49905; S22912

R:Howard, S.P.; Critch, J.; Bedi, A.

J. Bacteriol. 175, 6695-6703, 1993

A:Title: Isolation and analysis of eight exe genes and their involvement in extracellular

A:Reference number: A49905; MUID:94012544

A:Accession: B49905

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-119 <HOW>

A:Cross-references: GB:X66504; NID:g551420; PIDN:CAA47130.1; PID:g38830

C:Genetics:

A:Gene: exeI

C:Superfamily: secretion protein xcpv

Query Match 45.5%; Score 5; DB 2; Length 119;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TAPKA 8  
|||||  
DB 102 TAPKA 106

## RESULT 21

B81363

Hit-family protein Cj0898 [imported] - Campylobacter jejuni (strain NCTC 11168)

C:Species: Campylobacter jejuni

C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 18-Aug-2000

C:Accession: B81363

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chil  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba  
Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals

A:Reference number: A81250; MUID:20150912

A:Accession: B81363

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-121 <PAR>

A:Cross-references: GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB73156.1; PID:g696

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj0898

C:Superfamily: protein kinase C inhibitor; histidine triad homology

Query Match 45.5%; Score 5; DB 2; Length 121;  
Best Local Similarity 100.0%; Pred. No. 73;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APKAP 9  
|||||  
DB 33 APKAP 37

## RESULT 22

S49484

histone H2B - sea cucumber (Holothuria tubulosa)

C:Species: Holothuria tubulosa

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 23-Jul-1999

C:Accession: S49484

R:Drabant, B.; Louroutziatis, A.; Pratz, E.; Cornudella, L.; Doenecke, D.

submitted to the EMBL Data Library, September 1994

A:Description: Structure of histone H2B and H4 genes of the sea cucumber Holothuria t

A:Reference number: S49484

A:Accession: S49484

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-123 <DRA>

A:Cross-references: EMBL:246225; NID:g559456; PIDN:CAA86297.1; PID:g559457

C:Superfamily: histone H2B

Query Match 45.5%; Score 5; DB 2; Length 123;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APKAP 9  
|||||  
DB 2 APKAP 6

## RESULT 23

S09877

hypothetical protein UL110 - human cytomegalovirus (strain AD169)

C:Species: human cytomegalovirus, human herpesvirus 5

A:Note: host Homo sapiens (man)

C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 28-Jul-2000

C:Accession: S09877

R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohnl, R.; Brown, C.M.; Cerny, R.; Horsnell, T

M.; Barrall, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovi

A:Reference number: S09745; MUID:90269039

A:Accession: S09877

A>Status: nucleic acid sequence not shown; translation not shown

A:Reference number: A59233  
A:Accession: A59233  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-2121 <AND>  
A:Cross-references: GB:AF233269; PIDN:AAF34810.1  
C:Genetics:  
A:Gene: Myo28B1  
A:Cross-references: FlyBase:FBgn0040299  
A:Map position: 2, 28B1  
C:Superfamily: myosin motor domain homology  
F:69-725/Domain: myosin motor domain homology <MMO>

Query Match 54.5%; Score 6; DB 2; Length 2121;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 APKAP 10  
|||||  
Db 915 APKAP 920

## RESULT 14

S07073  
arabinogalactan protein - Italian ryegrass (fragments)  
C:Species: Lolium multiflorum (Italian ryegrass)  
C:Date: 21-Nov-1993 #sequence\_revision 01-Dec-1995 #text\_change 05-Dec-1998  
C:Accession: S07073  
R:Gleeson, P.A.; McNamara, M.; Wattenhall, R.E.H.; Stone, B.A.; Fincher, G.B.  
Biochem. J. 264, 857-862, 1989  
A:Title: Characterization of the hydroxyproline-rich protein core of an arabinogalactan-  
A:Reference number: S07073; MUID:90147544  
A:Accession: S07073  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-46 <GLE>  
A:Note: 19-His and 23-Leu were also found  
C:Keywords: hydroxyproline  
F:6,7,9,11,14,16,18,20,26,30,34,36,39/Modified site: hydroxyproline (Pro) #status experi

Query Match 45.5%; Score 5; DB 2; Length 46;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 APKAP 9  
|||||  
Db 10 APKAP 14

## RESULT 15

I47395  
histone H1 I-1 (clone T31) - midge (Chironomus thummi thummi) (fragment)  
C:Species: Chironomus thummi thummi  
C:Date: 24-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 03-May-1996  
C:Accession: I47395  
R:Schulze, E.; Trieschmann, L.; Schulze, B.; Schmidt, E.R.; Pitzel, S.; Zechel, K.; Gros  
Proc. Natl. Acad. Sci. U.S.A. 90, 2481-2485, 1993  
A:Title: Structural and functional differences between histone H1 sequence variants with  
A:Reference number: A47395; MUID:93211985  
A:Accession: I47395  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-87 <SCH>  
A:Note: sequence extracted from NCBI backbone (NCBIP:128535)  
C:Superfamily: histone H1

Query Match 45.5%; Score 5; DB 2; Length 67;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 APKAP 9  
|||||  
Db 27 APKAP 31

## RESULT 16

A47395  
histone H1 I-1 (N-terminal, clone L1234, globular domain) - midge (Chironomus thummi)  
C:Species: Chironomus thummi  
C:Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 03-May-1996  
C:Accession: A47395  
R:Schulze, E.; Trieschmann, L.; Schulze, B.; Schmidt, E.R.; Pitzel, S.; Zechel, K.; G  
Proc. Natl. Acad. Sci. U.S.A. 90, 2481-2485, 1993  
A:Title: Structural and functional differences between histone H1 sequence variants w  
A:Reference number: A47395; MUID:93211985  
A:Accession: A47395  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid; protein  
A:Residues: 1-70 <SCH>  
A:Experimental source: subsp. piger  
A:Note: sequence extracted from NCBI backbone (NCBIP:128533)  
C:Superfamily: histone H1

Query Match 45.5%; Score 5; DB 2; Length 70;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 APKAP 9  
|||||  
Db 27 APKAP 31

## RESULT 17

A44465  
sodium ion pump oxaloacetate decarboxylase subunit gamma - Salmonella typhimurium  
C:Species: Salmonella typhimurium  
C:Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 16-Feb-1997  
C:Accession: A44465  
R:Woehlke, G.; Wifling, K.; Dimroth, P.  
J. Biol. Chem. 267, 22798-22803, 1992  
A:Title: Sequence of the sodium ion pump oxaloacetate decarboxylase from Salmonella t  
A:Reference number: A44465; MUID:93054591  
A:Accession: A44465  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-83 <WOE>  
A:Experimental source: LT2  
A:Note: sequence extracted from NCBI backbone (NCBIP:118071)  
C:Keywords: transmembrane protein

Query Match 45.5%; Score 5; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 APKAP 9  
|||||  
Db 50 APKAP 54

## RESULT 18

S36634  
Ptl2 protein - Antirrhinum sp.  
C:Species: Antirrhinum sp.  
C:Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 05-Nov-1999  
C:Accession: S36634  
R:Baldwin, T.C.; Coen, E.S.; Dickinson, H.G.  
submitted to the EMBL Data Library, January 1993  
A:Reference number: S36634  
A:Accession: S36634  
A:Molecule type: mRNA  
A:Residues: 1-95 <BAL>

C:Superfamily: translation initiation factor IF-2; translation elongation factor Tu homolog  
 C:Keywords: GTP binding; nucleotide binding; P-loop  
 F:496-609/Domain: translation elongation factor Tu homology <ETU>  
 F:502-609/Region: nucleotide-binding motif A (P-loop)  
 F:606-609/Region: GTP-binding NKXD motif  
 F:642-644/Region: GTP-binding SAK/L motif  
 F:508,509,529,606,607,609,642/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #

Query Match 54.5%; Score 6; DB 2; Length 1001;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 APKAPT 10  
 |||||  
 Db 111 APKAPT 116

## RESULT 10

T14343  
 zinc finger RNA binding protein, chromosome-associated - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T14343  
 R:Meagher, M.J.; Schumacher, J.M.; Lee, K.; Holdcraft, R.W.; Edelhoff, S.; Distech, C.  
 Gene 228, 197-211, 1999  
 A:Title: Identification of ZFR, an ancient and highly conserved murine chromosome-associated  
 A:Reference number: 217994; MUID:99173884  
 A:Accession: T14343

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA

A:Residues: 1-1052 <MEA>  
 A:Cross-references: EMBL:AF071059; NID:g3293536; PID:g3293537; PIDN:AAC25762.1  
 A:Experimental source: testis

C:Genetics:  
 A:Gene: Zfr

A:Map position: 15A  
 C:Keywords: zinc finger

Query Match 54.5%; Score 6; DB 2; Length 1052;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 QTAPKA 8  
 |||||  
 Db 165 QTAPKA 170

## RESULT 11

RRNGSM  
 RNA-directed RNA polymerase (EC 2.7.7.48) - strawberry mild yellow edge-associated virus  
 N:Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase  
 C:Species: strawberry mild yellow edge-associated virus  
 C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 19-Jan-2001  
 C:Accession: JQ1426  
 R:Jelkmann, W.; Maiss, E.; Martin, R.R.  
 J. Gen. Virol. 73, 475-479, 1992

A:Title: The nucleotide sequence and genome organization of strawberry mild yellow edge-  
 A:Reference number: JQ1426; MUID:92166762  
 A:Accession: JQ1426

A:Molecule type: genomic RNA  
 A:Residues: 1-1323 <JEL>

A:Cross-references: GB:D12517; DBJ:D01227; NID:g222631; PIDN:BAA02082.1; PID:g222632

C:Superfamily: eggplant mosaic virus RNA-directed RNA polymerase  
 C:Keywords: ATP; nucleotide binding; nucleotidyltransferase; P-loop; RNA biosynthesis; P  
 F:605-612/Region: nucleotide-binding motif A (P-loop)  
 F:668-673/Region: nucleotide-binding motif B  
 F:611/Binding site: ATP (Lys) #status predicted

Query Match 54.5%; Score 6; DB 1; Length 1323;  
 Best Local Similarity 100.0%; Pred. No. 44;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 APKAPT 10  
 |||||  
 Db 463 APKAPT 468

## RESULT 12

T43214  
 ovrl protein - nematode (Onchocerca volvulus)  
 N:Alternate names: myosin-like antigen  
 C:Species: Onchocerca volvulus  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 20-Jun-2000  
 C:Accession: T43214; A44939; A54513; S27825  
 R:Triteeraprapab, S.; Richie, T.L.; Tuan, R.S.; Shepley, K.J.; Dinman, J.D.; Neubert,  
 Mol. Biochem. Parasitol. 69, 161-171, 1995  
 A:Title: Molecular cloning of a gene expressed during early embryonic development in  
 A:Reference number: 222341; MUID:95287898  
 A:Accession: T43214

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2022 <TRI>

A:Cross-references: EMBL:U12681; NID:g530824; PID:g530825; PIDN:AAA80009.1

A:Experimental source: specific host Homo sapiens  
 R:Erondou, N.E.; Donelson, J.E.

Mol. Biochem. Parasitol. 40, 213-224, 1990

A:Title: Characterization of a myosin-like antigen from Onchocerca volvulus.

A:Reference number: A44939; MUID:90301142

A:Accession: A44939

A:Molecule type: mRNA

A:Residues: 733-874, 'E', 876-916, 'S', 918-1038, 1040-1047, 'S', 1049-1283, 'E', 1285-1363 <E  
 A:Cross-references: GB:M30398

A:Note: the sequence is revised in GenBank entry ONGANTML, release 115, (PIDN:AAA2941  
 R:Donelson, J.E.; Duke, B.O.L.; Moser, D.; Zeng, W.; Erondou, N.E.; Lucius, R.; Renz,

Mol. Biochem. Parasitol. 31, 241-250, 1988

A:Title: Construction of Onchocerca volvulus cDNA libraries and partial characterization  
 A:Reference number: A54513; MUID:89127417

A:Accession: A54513

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 733-866 <DON>  
 A:Cross-references: GB:J03995; NID:g159874; PIDN:AAA29412.1; PID:g159875

R:Ngozi, E.; Erondou, N.E.; Donelson, J.E.

submitted to the EMBL Data Library, April 1990

A:Description: Characterization of a myosin-like antigen from Onchocerca volvulus.

A:Reference number: S27825

A:Accession: S27825

A:Molecule type: mRNA

A:Residues: 733-874, 'E', 876-916, 'S', 918-1038, 1040-1047, 'S', 1049-1283, 'E', 1285-1347, 'E  
 A:Cross-references: EMBL:M30398; NID:g159876; PIDN:AAA29413.1; PID:g159877

C:Genetics:

A:Gene: ovrl

C:Keywords: leucine zipper

Query Match 54.5%; Score 6; DB 2; Length 2022;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOOTAP 6  
 |||||  
 Db 302 QOOTAP 307

## RESULT 13

AS9233  
 myosin VII-like protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 17-Nov-2000

C:Accession: A59233

R:Anderson, J.B.; Yamashita, R.A.; Sellers, J.R.

submitted to GenBank, February 2000

A:Description: Complete cDNA for an unconventional Myosin (Class VII) in Drosophila m





```

979 4 36.4 220 2 JCS954 ribosomal protein
980 4 36.4 220 2 T18982 hypothetical prote
981 4 36.4 221 2 S49482 histone H1 - mouse
982 4 36.4 221 2 S33219 histone H1.C - Afr
983 4 36.4 221 2 S58685 hypothetical prote
984 4 36.4 221 2 T45044 hypothetical prote
985 4 36.4 221 2 T36514 hypothetical prote
986 4 36.4 221 2 T34554 hypothetical prote
987 4 36.4 222 2 JC1384 beta-casein precu
988 4 36.4 222 2 A32979 beta-casein precu
989 4 36.4 222 2 T16731 hypothetical prote
990 4 36.4 222 2 S66832 hypothetical prote
991 4 36.4 223 1 TJBPI.L tail assembly prot
992 4 36.4 223 2 A96622 conserved hypotet
993 4 36.4 223 2 T36952 conserved hypotet
994 4 36.4 224 1 KBR0A2 beta-casein precu
995 4 36.4 224 2 C86571 L4 ribosomal prote
996 4 36.4 224 2 H72055 ribosomal protein
997 4 36.4 224 2 T24802 hypothetical prote
998 4 36.4 224 2 B70327 ABC transporter -
999 4 36.4 225 2 F85718 probable tail comp
1000 4 36.4 225 2 C85475 hypothetical prote

```

## ALIGNMENTS

```

RESULT 1
A41487
protein p60 precursor - Listeria monocytogenes
N:Alternate names: Invasion-associated protein
C:Species: Listeria monocytogenes
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1996
C:Accession: A41487 B41487
R:Koehler, S.; Delmester-Waechter, M.; Chakraborty, T.; Lottspeich, F.; Goebel, W.
Infect. Immun. 58, 1943-1950, 1990
A:Title: The gene coding for protein p60 of Listeria monocytogenes and its use as a spec
A:Reference number: A41487; PMID:90256283
A:Accession: A41487
A:Molecule type: DNA
A:Residues: 1484 <KOE>
A:Cross-references: GB:X52268
A:Accession: B41487
A:Molecule type: protein
A:Residues: 28-49 <KO2>
C:Genetics:
A:Gene: lap
F:1.27/Domain: signal sequence #status predicted <SIG>
F:28-484/Product: protein 60 #status predicted <MAT>

```

```

Query Match 100.0%; Score 11; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 6; 9e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QOQTAPKAPTE 11
|||||
Db 291 QOQTAPKAPTE 301

```

```

RESULT 2
T29819
hypothetical protein F25E2.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T29819
R:Minx, P.; Le, T.
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of C. elegans cosmid F25E2.
A:Reference number: Z20691
A:Accession: T29819
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA

```

```

A:Residues: 1-343 <MIN>
A:Cross-references: EMBL:U50197; PIDD:AAA91256.1; CESP:F25E2.3
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F25E2.3
A:Introns: 9/3; 48/2; 66/3; 96/1; 176/3; 245/1; 311/1
C:Superfamily: acyl-CoA thioesterase II

```

```

Query Match 54.5%; Score 6; DB 2; Length 343;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 6 PKAPTE 11
|||||
Db 4 PKAPTE 9

```

```

RESULT 3
T22894
hypothetical protein F58B3.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22894
R:Harris, B.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19633
A:Accession: T22894
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-371 <NLL>
A:Cross-references: EMBL:Z73427; PIDD:CAA97799.1; GSPDB:GN00022; CESP:F58B3.7
A:Experimental source: clone F58B3
C:Genetics:
A:Gene: CESP:F58B3.7
A:Map position: 4
A:Introns: 19/1; 46/3; 317/3

```

```

Query Match 54.5%; Score 6; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 5 APKAPT 10
|||||
Db 255 APKAPT 260

```

```

RESULT 4
T37217
probable ATP/GTP binding protein - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37217
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21615
A:Accession: T37217
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-606 <OLI>
A:Cross-references: EMBL:AL031514; PIDD:CAA20595.1; GSPDB:GN000070; SCODEB:SC2H4.01
C:Genetics:
A:Gene: SCODEB:SC2H4.01

```

```

Query Match 54.5%; Score 6; DB 2; Length 606;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 TAPKAP 9
|||||

```

833	4	36.4	195	2	S52221	blue protein - Rho	906	4	36.4	208	2	S14700	hypothetical 23k p
834	4	36.4	195	2	T03522	blue protein - Rho	907	4	36.4	208	2	T07732	tuberculosis-relat
835	4	36.4	195	2	B69227	ribonucleoside-dip	908	4	36.4	209	2	G69694	ribosomal protein
836	4	36.4	195	2	T39603	Bac-family molecu	909	4	36.4	209	2	T44383	ribosomal protein
837	4	36.4	196	2	D69278	3-dehydroquinat	910	4	36.4	209	2	A59068	beta-casein varian
838	4	36.4	196	2	S43436	B cell activation	911	4	36.4	209	2	I40307	outer surface prot
839	4	36.4	196	2	C86858	hypothetical prote	912	4	36.4	209	2	C70630	hypothetical prote
840	4	36.4	197	2	A53126	syndecan-4 precu	913	4	36.4	210	2	A47733	GTP-binding protei
841	4	36.4	197	2	C70519	hypothetical prote	914	4	36.4	210	2	S36297	T-cell receptor ga
842	4	36.4	198	1	B64842	trp repressor-bind	915	4	36.4	210	2	A64161	hypothetical prote
843	4	36.4	198	1	A35851	nucd protein - Eme	916	4	36.4	210	2	T19239	hypothetical prote
844	4	36.4	198	2	G61675	riboflavin synthas	917	4	36.4	210	2	T28002	hypothetical prote
845	4	36.4	198	2	S45188	nicotinate-nucleot	918	4	36.4	211	2	T43338	adenovirus E1b int
846	4	36.4	198	2	JC5613	ryudocan precursor	919	4	36.4	211	2	S34729	GTP-binding protei
847	4	36.4	198	2	JC1457	remorin - potato	920	4	36.4	211	2	G75533	ribosomal protein
848	4	36.4	198	2	T07280	cysteine string pr	921	4	36.4	211	2	A42261	heparan sulfate pr
849	4	36.4	198	2	I52655	hypothetical prote	922	4	36.4	211	2	D65011	hypothetical prote
850	4	36.4	198	2	S70515	hypothetical prote	923	4	36.4	211	2	A65880	hypothetical prote
851	4	36.4	198	2	PQ0034	hypothetical prote	924	4	36.4	211	2	T10072	periplasmic protei
852	4	36.4	198	2	S48994	hypothetical prote	925	4	36.4	211	2	E81382	hypothetical prote
853	4	36.4	198	2	T19385	hypothetical prote	926	4	36.4	211	2	S28046	TUB8 protein - pot
854	4	36.4	199	2	J50430	prolactin - elepha	927	4	36.4	212	2	T51629	endoplasmatic reti
855	4	36.4	199	2	A82360	conserved hypotet	928	4	36.4	212	2	A83521	conserved hypotet
856	4	36.4	199	2	S23379	desiccation stress	929	4	36.4	213	1	HSRB13	histone H1.3 - rab
857	4	36.4	199	2	A82721	conserved hypotet	930	4	36.4	213	1	S24363	ribosomal protein
858	4	36.4	199	2	S15164	hypothetical prote	931	4	36.4	213	2	JE0247	Ig lambda chain NI
859	4	36.4	199	2	S72728	probable L-a-amino	932	4	36.4	213	2	S21066	Ig lambda chain V
860	4	36.4	199	2	S05509	photosystem II oxy	933	4	36.4	213	2	S75886	hypothetical prote
861	4	36.4	199	2	S77228	hypothetical prote	934	4	36.4	213	2	E83979	hypothetical prote
862	4	36.4	199	2	T18879	hypothetical prote	935	4	36.4	213	2	G65673	hypothetical prote
863	4	36.4	199	2	T29151	hypothetical prote	936	4	36.4	213	2	F64854	ycfM protein precu
864	4	36.4	199	2	D83773	hypothetical prote	937	4	36.4	213	2	F70946	ribosomal regulato
865	4	36.4	200	1	MORT3V	mosesin alkali ligh	938	4	36.4	213	2	E82698	conserved hypotet
866	4	36.4	200	2	G75550	conserved hypotet	939	4	36.4	214	1	JC4808	ribosomal protein
867	4	36.4	201	2	UC5337	GTP-binding protei	940	4	36.4	214	2	C82255	adenylylate kinase V
868	4	36.4	201	2	D70619	probable infc prot	941	4	36.4	214	2	S39831	hypothetical prote
869	4	36.4	201	2	S37847	hypothetical prote	942	4	36.4	214	2	S17807	hypothetical prote
870	4	36.4	201	2	A47742	Rho-GDP-dissociat	943	4	36.4	214	2	H85359	hypothetical prote
871	4	36.4	202	2	T46523	probable cyclase-d	944	4	36.4	215	2	C85584	probable tail comp
872	4	36.4	202	2	I53137	fibroglycan (synde	945	4	36.4	215	2	A86437	hypothetical prote
873	4	36.4	202	2	A42410	ryudocan precursor	946	4	36.4	215	2	H82483	probable hexulose-
874	4	36.4	202	2	S06882	vitC2 protein - Ag	947	4	36.4	215	2	S73929	ribosomal protein
875	4	36.4	202	2	D83578	conserved hypotet	948	4	36.4	215	2	H86774	hypothetical prote
876	4	36.4	202	2	T48709	hypothetical prote	949	4	36.4	215	2	B70622	probable ribosomal
877	4	36.4	202	2	S64051	hypothetical prote	950	4	36.4	215	2	T22446	hypothetical prote
878	4	36.4	203	1	T31684	probable bacitraci	951	4	36.4	215	2	PC4355	connectin - chicke
879	4	36.4	203	2	H65052	hydrogenase (EC 1.	952	4	36.4	216	2	JE0245	Ig Lambda chain NI
880	4	36.4	203	2	T12808	hypothetical prote	953	4	36.4	216	2	A42193	Ig Lambda chain (B
881	4	36.4	203	2	H86713	hypothetical prote	954	4	36.4	216	2	S03401	Ig lambda chain (K
882	4	36.4	203	2	F96694	hypothetical prote	955	4	36.4	216	2	S29258	Ig lambda chain V
883	4	36.4	203	2	C85921	hypothetical prote	956	4	36.4	216	2	S69130	Ig lambda chain (D
884	4	36.4	203	2	B83166	hypothetical prote	957	4	36.4	216	2	S36288	T-cell receptor ga
885	4	36.4	203	2	T18679	hypothetical prote	958	4	36.4	216	2	T33964	hypothetical prote
886	4	36.4	204	1	D70111	H+-transporting AT	959	4	36.4	216	2	T00121	hypothetical prote
887	4	36.4	204	2	G75307	transcription regu	960	4	36.4	217	2	JE0246	Ig lambda chain NI
888	4	36.4	204	2	S44829	F54F2. 4 protein -	961	4	36.4	217	2	T37859	probable transcrip
889	4	36.4	204	2	G82468	MutR/nudix family	962	4	36.4	217	2	T30805	hypothetical prote
890	4	36.4	204	2	A82216	hypothetical prote	963	4	36.4	217	2	F82612	hypothetical prote
891	4	36.4	205	2	T32959	hypothetical prote	964	4	36.4	217	2	S45446	hypothetical prote
892	4	36.4	205	2	T02662	hml protein homol	965	4	36.4	218	2	S34359	ornithine carbamoy
893	4	36.4	205	2	B82237	conserved hypotet	966	4	36.4	218	2	G70550	butyrate--acetace
894	4	36.4	205	2	H85072	hypothetical prote	967	4	36.4	218	2	A71304	hypothetical prote
895	4	36.4	206	1	HSTR1R	histone H1 - rainb	968	4	36.4	218	2	B72116	hypothetical prote
896	4	36.4	206	2	G70307	adenylylate kinase (	969	4	36.4	218	2	T30000	hypothetical prote
897	4	36.4	206	2	UC5491	ABC transporter Am	970	4	36.4	219	1	HSXLIB	histone H1B - Afri
898	4	36.4	206	2	B81819	hypothetical prote	971	4	36.4	219	2	T08083	H+-transporting AT
899	4	36.4	206	2	T34979	probable lipoprote	972	4	36.4	219	2	E75025	1-Isospartyl prot
900	4	36.4	206	2	G86279	hypothetical prote	973	4	36.4	219	2	T26632	hypothetical prote
901	4	36.4	206	2	T32854	hypothetical prote	974	4	36.4	219	2	A86649	hypothetical prote
902	4	36.4	207	2	F69609	cytochrome-c oxida	975	4	36.4	219	2	S42674	adhesive protein -
903	4	36.4	208	2	B49444	Ig lambda chain (N	976	4	36.4	220	2	D83390	cadonate denydrat
904	4	36.4	208	2	I70195	histone H1t - rhos	977	4	36.4	220	2	I51447	histone H1B - Afri
905	4	36.4	208	2	B86887	50S ribosomal prot	978	4	36.4	220	2	A46597	acidic calmodulin-

687	4	36.4	160	2	S18658	variant surface an	760	4	36.4	179	2	H85429	ribosomal protein
688	4	36.4	160	2	H69015	ribosomal protein	761	4	36.4	179	2	T48372	hypothetical prote
689	4	36.4	160	2	F84350	hypothetical prote	762	4	36.4	180	1	J00986	lipid body-associ
690	4	36.4	160	2	A82733	conserved hypotet	763	4	36.4	180	2	S72923	hypothetical prote
691	4	36.4	161	2	J00352	heat shock protein	764	4	36.4	180	2	F83797	hypothetical prote
692	4	36.4	161	2	A29557	calcium-binding pr	765	4	36.4	180	2	G83985	hypothetical prote
693	4	36.4	161	2	S05547	dehydrin 17 - barl	766	4	36.4	181	2	D85636	hypothetical prote
694	4	36.4	161	2	B84366	hypothetical prote	767	4	36.4	182	2	E72630	hypothetical prote
695	4	36.4	162	2	JC7157	calcium vector pro	768	4	36.4	182	2	F69387	hypothetical prote
696	4	36.4	162	2	T50504	DRI-like protein -	769	4	36.4	182	2	T23828	hypothetical prote
697	4	36.4	162	2	S28122	gas-vesicle operon	770	4	36.4	183	2	F82882	ribosome releasing
698	4	36.4	162	2	T40004	hypothetical prote	771	4	36.4	183	2	S59091	conserved hypotet
699	4	36.4	162	2	S74418	hypothetical prote	772	4	36.4	183	2	E69388	hypothetical prote
700	4	36.4	163	2	F83059	acetylactate synth	773	4	36.4	183	2	E96582	hypothetical prote
701	4	36.4	163	2	H81066	acetylactate synth	774	4	36.4	183	2	B64547	hypothetical prote
702	4	36.4	163	2	F49905	protein secretion	775	4	36.4	183	2	T28711	hypothetical prote
703	4	36.4	163	2	B44813	probable molybdenu	776	4	36.4	184	2	A82013	dtDP-4-dehydrotham
704	4	36.4	163	2	E75416	hypothetical prote	777	4	36.4	184	2	A45281	transcription fact
705	4	36.4	163	2	T33130	hypothetical prote	778	4	36.4	184	2	H72070	heat shock protein
706	4	36.4	164	1	S74345	hypothetical prote	779	4	36.4	184	2	B86553	HSP-70 cofactor [1
707	4	36.4	164	2	T37174	protein-tyrosine-p	780	4	36.4	184	2	A69525	hypothetical prote
708	4	36.4	164	2	D85928	hypothetical prote	781	4	36.4	184	2	E96594	unknown protein, 6
709	4	36.4	165	1	S70213	hypothetical prote	782	4	36.4	184	2	A82631	hypothetical prote
710	4	36.4	165	2	D96758	hypothetical prote	783	4	36.4	184	2	S73868	lipoprotein signal
711	4	36.4	165	2	A60611	18K microfilariat	784	4	36.4	184	2	S25815	hypothetical prote
712	4	36.4	165	2	S55766	transcription init	785	4	36.4	185	2	S59560	histone H1.41 - ga
713	4	36.4	165	2	H83918	hypothetical prote	786	4	36.4	185	2	S38959	ribosomal protein
714	4	36.4	166	2	B72626	hypothetical prote	787	4	36.4	185	2	H82799	fibrillinl1 XF0487
715	4	36.4	167	1	S37878	IgE-dependent hist	788	4	36.4	185	2	D84456	hypothetical prote
716	4	36.4	167	2	S50808	hypothetical prote	789	4	36.4	186	1	R5B78E	ribosomal protein
717	4	36.4	167	2	E83743	4-hydroxybenzoyl-C	790	4	36.4	187	2	S77483	adenylate kinase (
718	4	36.4	167	2	H72579	hypothetical prote	791	4	36.4	188	2	S25979	ribosomal protein
719	4	36.4	168	2	JN0073	glycophorin A - mo	792	4	36.4	188	2	T26939	hypothetical prote
720	4	36.4	168	2	E75257	molybdenum cofacto	793	4	36.4	188	2	E65207	hypothetical prote
721	4	36.4	168	2	T02036	early light-induce	794	4	36.4	188	2	C86092	hypothetical prote
722	4	36.4	168	2	S65242	hypothetical prote	795	4	36.4	188	2	D82240	conserved hypotet
723	4	36.4	169	2	T20606	hypothetical prote	796	4	36.4	189	2	S71465	fibronectin - chic
724	4	36.4	169	2	D81180	16S rRNA processin	797	4	36.4	189	2	G75349	conserved hypotet
725	4	36.4	169	2	B81924	probable 16S rRNA	798	4	36.4	189	2	T43766	hypothetical prote
726	4	36.4	169	2	S47066	finger protein HZF	799	4	36.4	189	2	T21254	hypothetical prote
727	4	36.4	170	2	E81662	Holliday junction	800	4	36.4	190	2	A82015	dtDP-4-dehydrotham
728	4	36.4	170	2	JT0588	hypothetical 20K p	801	4	36.4	190	2	S25740	Ig lambda chain -
729	4	36.4	171	2	E83140	phosphatidylglycer	802	4	36.4	190	2	S66981	ribosomal protein
730	4	36.4	171	2	H81029	lipoprotein NMB189	803	4	36.4	190	2	E64775	probable lipoprote
731	4	36.4	171	2	E81974	lipoprotein NMA055	804	4	36.4	190	2	G85542	glycoprotein/polys
732	4	36.4	171	2	S40769	hypothetical prote	805	4	36.4	190	2	T43013	conserved hypotet
733	4	36.4	172	1	F1B83F	translation initia	806	4	36.4	190	2	T25462	hypothetical prote
734	4	36.4	172	2	B64320	acetylactate synth	807	4	36.4	191	2	S56012	hypothetical prote
735	4	36.4	172	2	S77842	probable asparagin	808	4	36.4	191	2	S40177	XS-2 protein (homo
736	4	36.4	172	2	A45637	microcolte surface	809	4	36.4	191	2	H81444	probable cytochrom
737	4	36.4	172	2	F82309	autoinducer-2 prod	810	4	36.4	191	2	T35508	hypothetical prote
738	4	36.4	172	2	T36713	hypothetical prote	811	4	36.4	192	2	S42963	viral infectivity
739	4	36.4	172	2	I40349	ribosomal protein	812	4	36.4	192	2	T09382	vif protein - huma
740	4	36.4	173	2	S69519	hypothetical prote	813	4	36.4	192	2	A48353	genome polyprotein
741	4	36.4	174	2	H83481	hypothetical prote	814	4	36.4	192	2	B48353	genome polyprotein
742	4	36.4	174	2	T51542	hypothetical prote	815	4	36.4	192	2	A83798	transcriptional re
743	4	36.4	175	2	E83777	O6-methylguanine D	816	4	36.4	192	2	F72805	gp49 protein - Myc
744	4	36.4	175	2	S70915	major oleosin nap-	817	4	36.4	193	2	B69178	hypothetical prote
745	4	36.4	175	2	B86767	hypothetical prote	818	4	36.4	193	2	T44106	hypothetical prote
746	4	36.4	176	2	G70658	probable aroK prot	819	4	36.4	193	2	G64356	hypothetical prote
747	4	36.4	176	2	A73624	hypothetical prote	820	4	36.4	194	1	C75206	L-fuculose-phospha
748	4	36.4	176	2	T35459	hypothetical prote	821	4	36.4	194	1	HSTR1	histone H1 - trout
749	4	36.4	177	2	F84782	60S ribosomal prot	822	4	36.4	194	1	HSN10	conserved hypotet
750	4	36.4	177	2	S53952	proteolase 2 precu	823	4	36.4	194	2	A82072	hypothetical prote
751	4	36.4	177	2	D82638	hypothetical prote	824	4	36.4	194	2	E69899	hypothetical prote
752	4	36.4	177	2	D83135	hypothetical prote	825	4	36.4	194	2	S72877	transcription regu
753	4	36.4	178	1	A39070	endothelin 2 precu	826	4	36.4	194	2	T44862	hypothetical prote
754	4	36.4	178	2	S74564	hypothetical prote	827	4	36.4	194	2	B64159	hypothetical prote
755	4	36.4	178	2	T49712	conserved hypotet	828	4	36.4	194	2	T14746	myosin alkali 11ph
756	4	36.4	178	2	E75340	hypothetical prote	829	4	36.4	195	1	MOHU3V	coagulogen precurs
757	4	36.4	179	2	H85816	hypothetical prote	830	4	36.4	195	1	WCNCA	BHV-1 protein homo
758	4	36.4	179	2	F71854	peptidoglycan asso	831	4	36.4	195	2	S24228	hypothetical prote
759	4	36.4	179	2	E64660		832	4	36.4	195	2	E70555	

541	4	36.4	130	2	S46702	hypothetical prote
542	4	36.4	131	1	RDEB15	funarate reductase
543	4	36.4	131	2	S56380	succinate dehydrog
544	4	36.4	131	2	E86111	hypothetical prote
545	4	36.4	131	2	E30552	T-cell activation
546	4	36.4	131	2	H75323	hypothetical prote
547	4	36.4	132	1	RWMSAV	T-cell receptor al
548	4	36.4	132	1	G69256	conserved hypoteth
549	4	36.4	132	2	T48278	hypothetical prote
550	4	36.4	133	2	S61662	dolichyl-diphospho
551	4	36.4	134	1	AWHUB	natriuretic peptid
552	4	36.4	134	2	G82411	conserved hypoteth
553	4	36.4	134	2	JC6091	kinetoplast DNA-as
554	4	36.4	134	2	D64744	exopolysaccharide
555	4	36.4	134	2	F85504	hypothetical prote
556	4	36.4	134	2	B72737	hypothetical prote
557	4	36.4	134	2	D81096	hypothetical prote
558	4	36.4	135	2	S56687	histone H2B153 - w
559	4	36.4	135	2	C70890	hypothetical prote
560	4	36.4	136	2	S56684	histone H2B-6 - wh
561	4	36.4	136	2	S76416	hypothetical prote
562	4	36.4	136	2	T07975	probable arabinoga
563	4	36.4	136	2	T07945	probable arabinoga
564	4	36.4	136	2	C72575	hypothetical prote
565	4	36.4	137	2	T02035	histone H2B - maiz
566	4	36.4	137	2	T22308	hypothetical prote
567	4	36.4	138	2	C27577	T-cell receptor al
568	4	36.4	138	2	S56685	histone H2B-8 - wh
569	4	36.4	138	2	B84793	probable histone H
570	4	36.4	138	2	T45905	histone H2B-like p
571	4	36.4	138	2	T39504	hypothetical prote
572	4	36.4	138	2	F86117	probable transposa
573	4	36.4	139	2	E65036	hypothetical prote
574	4	36.4	139	2	T06389	histone H2B-2 - to
575	4	36.4	139	2	G85904	probable chlorodex
576	4	36.4	140	2	PH0134	Ig Lambda chain pr
577	4	36.4	140	2	JC4607	hydriophobin 1 pre
578	4	36.4	140	2	C69269	hypothetical prote
579	4	36.4	140	2	G65090	hypothetical prote
580	4	36.4	140	2	H85963	probable enzyme yq
581	4	36.4	141	2	G69904	capsular polysacch
582	4	36.4	141	2	A27482	vasotocin / neurop
583	4	36.4	141	2	H70698	hypothetical prote
584	4	36.4	141	2	H75583	probable copper re
585	4	36.4	141	2	PC1294	triphosphite surfac
586	4	36.4	141	2	T46380	hypothetical prote
587	4	36.4	141	2	C72606	hypothetical prote
588	4	36.4	142	2	T06393	histone H2B - toma
589	4	36.4	142	2	T10078	hypothetical prote
590	4	36.4	142	2	G84201	diadenosine tetrap
591	4	36.4	143	1	GPYF-	leghemoglobin I -
592	4	36.4	143	2	I48406	histone H2A.X - mo
593	4	36.4	143	2	S07631	histone H2A.X - hu
594	4	36.4	143	2	C86061	hypothetical prote
595	4	36.4	143	2	S10208	hypothetical 16.5K
596	4	36.4	143	2	F86168	hypothetical prote
597	4	36.4	143	2	E30338	DNA transport mach
598	4	36.4	143	2	A83431	type III export pr
599	4	36.4	144	2	C69889	hypothetical prote
600	4	36.4	144	2	T25028	hypothetical prote
601	4	36.4	144	2	A86795	hypothetical prote
602	4	36.4	144	2	T01476	hypothetical prote
603	4	36.4	145	2	T08063	histone H2B - peep
604	4	36.4	145	2	H72544	hypothetical prote
605	4	36.4	145	2	F83834	hypothetical prote
606	4	36.4	145	2	T01424	pectinesterase hom
607	4	36.4	145	2	F82133	hypothetical prote
608	4	36.4	145	2	F96766	protein actin F2P9
609	4	36.4	146	2	T12137	leghemoglobin K -
610	4	36.4	146	2	S47771	hypothetical 17.1k
611	4	36.4	146	2	B64053	ferric uptake regu
612	4	36.4	146	2	T07125	plastid-lipid-asso
613	4	36.4	146	2	B83700	hypothetical prote
614	4	36.4	146	2	G86029	hypothetical prote
615	4	36.4	147	2	T18577	peptidylprolyl iso
616	4	36.4	147	2	T09722	histone H2B1 - upl
617	4	36.4	147	2	A70977	probable ribosomal
618	4	36.4	147	2	S72991	ribosomal protein
619	4	36.4	147	2	S00848	ribonectin, trans
620	4	36.4	147	2	T30616	hypothetical prote
621	4	36.4	147	2	G72626	probable grp cyclo
622	4	36.4	147	2	S32015	flagellum-associat
623	4	36.4	147	2	T13202	hypothetical prote
624	4	36.4	148	2	T37532	ubiquitin-conjugat
625	4	36.4	148	2	B45637	merozoite surfac
626	4	36.4	148	2	E81788	conserved hypoteth
627	4	36.4	148	2	T22508	hypothetical prote
628	4	36.4	149	2	S23626	Ig lambda chain V
629	4	36.4	149	2	D82178	hypothetical prote
630	4	36.4	150	1	WMVZR2	I7K protein - vacc
631	4	36.4	150	2	S28049	histone H2B - maiz
632	4	36.4	150	2	F72163	AIL protein - vari
633	4	36.4	150	2	C42517	AIL protein - vacc
634	4	36.4	150	2	B36848	AIL protein - vari
635	4	36.4	150	2	T28542	hypothetical prote
636	4	36.4	150	2	T37387	VLRF-2, late gene
637	4	36.4	150	2	T24956	hypothetical prote
638	4	36.4	150	2	T37513	probable acetyltra
639	4	36.4	150	2	S58172	mithramycin polyke
640	4	36.4	150	2	T52587	probable arabinoga
641	4	36.4	150	2	T48611	ag86 protein - Ara
642	4	36.4	150	2	E83057	hypothetical prote
643	4	36.4	151	1	T17314	hypothetical prote
644	4	36.4	151	1	MUBPA7	N-acetyluramoyl-L
645	4	36.4	151	2	S07506	L
646	4	36.4	151	2	S28048	histone H2B - maiz
647	4	36.4	151	2	D84688	probable histone H
648	4	36.4	151	2	E85178	high affinity ribo
649	4	36.4	152	1	DABP73	adenosylmethionine
650	4	36.4	152	2	S14236	Ig gamma-1 chain C
651	4	36.4	152	2	S22323	histone H2B - whea
652	4	36.4	152	2	T17873	hypothetical prote
653	4	36.4	153	2	T44263	tRNA-pseudouridine
654	4	36.4	153	2	A64674	hypothetical prote
655	4	36.4	153	2	F71841	hypothetical prote
656	4	36.4	153	2	D64364	formate hydrogeny
657	4	36.4	154	1	DFSECM	aspartate carbamoy
658	4	36.4	154	2	T02077	histone H2B - maiz
659	4	36.4	154	2	T08308	hypothetical prote
660	4	36.4	154	2	B83444	conserved hypoteth
661	4	36.4	154	2	E72485	hypothetical prote
662	4	36.4	155	2	T05169	Les1 protein homol
663	4	36.4	155	2	S56258	hypothetical prote
664	4	36.4	155	2	G83774	hypothetical prote
665	4	36.4	155	2	E69482	hypothetical prote
666	4	36.4	155	2	T16855	hypothetical prote
667	4	36.4	155	2	T00134	hypothetical prote
668	4	36.4	156	1	R3R73A	ribosomal protein
669	4	36.4	156	2	A81015	conserved hypoteth
670	4	36.4	156	2	G83199	conserved hypoteth
671	4	36.4	156	2	T02806	probable membrane
672	4	36.4	156	2	S74733	hypothetical prote
673	4	36.4	157	1	JC2444	ribosomal protein
674	4	36.4	157	2	JN0549	ribosomal protein
675	4	36.4	157	2	S57603	hypothetical prote
676	4	36.4	157	2	T23000	hypothetical prote
677	4	36.4	157	2	T50859	hypothetical prote
678	4	36.4	158	2	E84956	response regulator
679	4	36.4	158	2	D81820	acetylactate synth
680	4	36.4	158	2	C71406	hypothetical prote
681	4	36.4	158	2	S74730	hypothetical prote
682	4	36.4	159	2	T06768	disease resistance
683	4	36.4	159	2	H75391	hypothetical prote
684	4	36.4	159	2	J00136	hypothetical 17.8k
685	4	36.4	160	1	S17194	endothelin 2 precu
686	4	36.4	160	2	C69526	histidine triad pr

395	4	36.4	104	2	F85359	hypothetical prote
396	4	36.4	104	2	B39435	hypothetical prote
397	4	36.4	105	1	L2HU	Ig lambda chain C
398	4	36.4	105	1	ASLJGR2	vpr protein - huma
399	4	36.4	105	1	ASLJGR	vpr protein - huma
400	4	36.4	105	2	S12156	vpr protein - huma
401	4	36.4	105	2	S53095	hypothetical prote
402	4	36.4	105	2	E64550	apolipoprotein 1 p
403	4	36.4	105	2	A26602	probable membrane
404	4	36.4	105	2	S48425	hypothetical IS110
405	4	36.4	105	2	E81798	hypothetical prote
406	4	36.4	106	1	A70701	thioedoxin - Rhod
407	4	36.4	106	1	A35135	conserved hypochla
408	4	36.4	106	2	F69055	probable periplasm
409	4	36.4	106	2	H72547	hypothetical prote
410	4	36.4	106	2	E72461	protein disulfide-
411	4	36.4	107	2	T17951	subtilisin-trypsin
412	4	36.4	107	2	JX0361	Ig lambda chain V-
413	4	36.4	107	2	B46516	T-cell receptor V-
414	4	36.4	107	2	S40133	hypothetical prote
415	4	36.4	107	2	T31471	hypothetical prote
416	4	36.4	107	2	T27713	Ig kappa chain V-I
417	4	36.4	108	1	K1H0WE	thioedoxin [valid
418	4	36.4	109	1	TXEC	chlorodioxin - Salm
419	4	36.4	109	1	S35497	Ig lambda chain V-
420	4	36.4	109	2	L1HUEP	hypothetical prote
421	4	36.4	109	2	S77093	NADH dehydrogenase
422	4	36.4	110	2	S15107	Ig lambda chain V-
423	4	36.4	110	2	S57408	Ig light chain V-J
424	4	36.4	110	2	S57428	T-cell receptor al
425	4	36.4	110	2	A49056	hypothetical prote
426	4	36.4	110	2	T48830	hypothetical prote
427	4	36.4	110	2	E82745	50S ribosomal prot
428	4	36.4	110	2	T02713	Ig lambda chain V-
429	4	36.4	111	1	L1H0NM	Ig lambda chain V-
430	4	36.4	111	1	L1H0NG	Ig lambda chain V-
431	4	36.4	111	1	L2H0VL	Ig lambda chain V-
432	4	36.4	111	2	S47009	Ig lambda chain VI
433	4	36.4	111	2	S47185	Ig lambda chain -
434	4	36.4	111	2	S22899	T-cell receptor al
435	4	36.4	111	2	T06418	pathogenesis-relat
436	4	36.4	111	2	E56556	fork head homolog
437	4	36.4	112	1	L1H0HA	Ig lambda chain V-
438	4	36.4	112	1	L1H0HM	Ig lambda chain V-
439	4	36.4	112	2	D44151	Ig lambda chain V
440	4	36.4	112	2	S51148	antibody light cha
441	4	36.4	112	2	C44151	Ig lambda chain V
442	4	36.4	112	2	S03503	T-cell receptor al
443	4	36.4	112	2	S40284	protein-tyrosine-p
444	4	36.4	112	2	A46717	collipase precursor
445	4	36.4	112	2	F75628	hypothetical prote
446	4	36.4	112	2	A84369	hypothetical prote
447	4	36.4	112	2	G81802	insertion element
448	4	36.4	113	1	KVMS75	Ig kappa chain V r
449	4	36.4	113	2	S42571	subtilisin inhibit
450	4	36.4	113	2	T36670	probable anti-sigm
451	4	36.4	114	2	C86362	cytochrome C (lipo
452	4	36.4	114	2	E84971	hypothetical prote
453	4	36.4	115	2	T07233	ribosomal protein
454	4	36.4	115	2	T48602	hypothetical prote
455	4	36.4	115	2	D84501	hypothetical prote
456	4	36.4	116	2	S62681	ribosomal protein
457	4	36.4	116	2	S24989	ribosomal protein
458	4	36.4	116	2	B40617	transcription repr
459	4	36.4	116	2	E72509	probable non speci
460	4	36.4	117	2	S04525	Ig lambda chain pr
461	4	36.4	117	2	S23627	Ig lambda chain pr
462	4	36.4	117	2	B44253	alveolar macrophag
463	4	36.4	117	2	S31130	hypothetical prote
464	4	36.4	117	2	T30740	hypothetical prote
465	4	36.4	117	2	A83312	conserved hypotet
466	4	36.4	117	2	JH0658	histone H1-like pr
467	4	36.4	118	1	R7MCML	ribosomal protein
468	4	36.4	118	2	S47077	Wnt-B protein - F1
469	4	36.4	118	2	S12627	Ig lambda chain pr
470	4	36.4	118	2	S21918	T-cell receptor al
471	4	36.4	118	2	S12440	Ig lambda chain (M
472	4	36.4	118	2	S12441	Ig lambda chain (K
473	4	36.4	118	2	S12442	Ig lambda chain (K
474	4	36.4	118	2	T07177	probable oleoyl-la
475	4	36.4	118	2	S49913	cryptogenin - Phyto
476	4	36.4	118	2	S49905	acidic elicitin A1
477	4	36.4	118	2	G64544	hypothetical prote
478	4	36.4	118	2	F71964	hypothetical prote
479	4	36.4	118	2	E33958	hypothetical 14k p
480	4	36.4	118	2	B85686	unknown protein en
481	4	36.4	118	2	T05520	geranylgeranylated
482	4	36.4	119	1	PSOXA	phospholipase A2 (
483	4	36.4	119	2	T14396	lipid transfer pro
484	4	36.4	119	2	T07984	lipid transfer pro
485	4	36.4	119	2	T06007	ribosomal protein
486	4	36.4	119	2	S76276	hypothetical prote
487	4	36.4	119	2	E64940	hypothetical prote
488	4	36.4	119	2	F85790	hypothetical prote
489	4	36.4	119	2	B72526	hypothetical prote
490	4	36.4	119	2	B84451	probable nitrilase
491	4	36.4	120	2	A61301	histone H2B - Tetra
492	4	36.4	120	2	F86395	60S ribosomal prot
493	4	36.4	120	2	G65009	hypothetical prote
494	4	36.4	121	2	S56686	histone H2B123 - w
495	4	36.4	121	2	B82946	ribosomal protein
496	4	36.4	121	2	S11671	hypothetical prote
497	4	36.4	121	2	D71527	hypothetical prote
498	4	36.4	121	2	S75660	hypothetical prote
499	4	36.4	122	2	A27097	histone H2B.1 - Te
500	4	36.4	122	2	B27097	histone H2B.2 - Te
501	4	36.4	122	2	S77492	ribosomal protein
502	4	36.4	123	2	F72630	hypothetical prote
503	4	36.4	123	2	E64418	conserved hypotet
504	4	36.4	123	2	S49906	HAe (highly acidic
505	4	36.4	123	2	B75546	hypothetical prote
506	4	36.4	123	2	D86755	prophage p12 prote
507	4	36.4	123	2	A64525	hypothetical prote
508	4	36.4	124	2	I40348	ribosomal protein
509	4	36.4	124	2	I40350	ribosomal protein
510	4	36.4	124	2	E70148	ribosomal protein
511	4	36.4	124	2	S72681	probable membrane
512	4	36.4	124	2	S38476	lys protein - pha
513	4	36.4	125	1	F1SP4	photosystem I chal
514	4	36.4	125	1	H64637	probable translati
515	4	36.4	125	2	A81310	30S ribosomal prot
516	4	36.4	125	2	E70332	conserved hypotet
517	4	36.4	125	2	A71878	hypothetical prote
518	4	36.4	125	2	T43414	DNA-directed RNA p
519	4	36.4	125	2	S74966	hypothetical prote
520	4	36.4	125	2	T00706	hypothetical prote
521	4	36.4	125	2	C72471	hypothetical prote
522	4	36.4	126	2	S53545	probable membrane
523	4	36.4	126	2	S09919	hypothetical prote
524	4	36.4	126	2	S75837	hypothetical prote
525	4	36.4	127	2	C86064	thioedoxin 1 (lipo
526	4	36.4	127	2	T36833	ribosomal protein
527	4	36.4	127	2	C84502	En/Spm-like transp
528	4	36.4	127	2	F82405	glyoxylase I fam11
529	4	36.4	128	2	A60215	myelin basic prote
530	4	36.4	128	2	A72404	hypothetical prote
531	4	36.4	128	2	T40344	hypothetical zinc
532	4	36.4	128	2	PEO006	histidinol dehydro
533	4	36.4	129	1	K1H0DI	Ig kappa chain pre
534	4	36.4	129	2	S78058	Ig lambda chain pr
535	4	36.4	129	2	D49094	methylnalonyl-CoA
536	4	36.4	130	1	L1H0BL	Ig lambda chain pr
537	4	36.4	130	2	S09712	Ig lambda chain V
538	4	36.4	130	2	S78057	Ig lambda chain pr
539	4	36.4	130	2	A75010	hypothetical prote
540	4	36.4	130	2	E70353	hypothetical prote

249	5	45.5	1929	2	T21559	hypothetical prote
250	5	45.5	1994	2	D86452	protein F6N18.13 [
251	5	45.5	2187	2	T30826	nascent polypeptid
252	5	45.5	2212	2	T28157	erythrocyte membra
253	5	45.5	2278	1	S56274	FAB1 protein - yea
254	5	45.5	2647	1	T28161	hypothetical prote
255	5	45.5	2756	2	T30183	hypothetical prote
256	5	45.5	6805	2	S20901	titlin - rabbit (fir
257	5	45.5	18344	1	I38344	titlin, cardiac mus
258	4	36.4	26926	9	PC7078	unidentified 48.7k
259	4	36.4	9	2	PN0115	insulin-like growt
260	4	36.4	20	2	E41299	T-cell receptor al
261	4	36.4	21	2	D47256	kinetoplast DNA-as
262	4	36.4	22	2	PC1272	subtilisin inhibit
263	4	36.4	34	2	S65417	pyruvate synthase
264	4	36.4	35	2	S69599	potassium channel
265	4	36.4	36	2	A82208	hypothetical prote
266	4	36.4	38	2	S23173	photosystem I chnl
267	4	36.4	40	2	S32045	Ig gamma-1 chain C
268	4	36.4	42	2	T36238	hypothetical prote
269	4	36.4	45	1	HPBO	haptoglobin precu
270	4	36.4	47	2	FA7395	histone H1 II-1 (c
271	4	36.4	47	2	S43964	hypothetical prote
272	4	36.4	48	2	S74261	tiopronin T3f, fast
273	4	36.4	51	2	I39691	cytochrome-c3 hydr
274	4	36.4	51	2	B45246	orf 3' of ldh - lac
275	4	36.4	52	2	H47395	histone H1 IIf-1 (
276	4	36.4	52	2	S01945	myosin catalytic I
277	4	36.4	53	2	F47395	histone H1 II-1 (c
278	4	36.4	53	2	B47395	histone H1 II-1 (c
279	4	36.4	53	2	C47395	histone H1 II-1 (c
280	4	36.4	53	2	D47395	histone H1 II-1 (c
281	4	36.4	53	2	G47395	histone H1 II-2 (c
282	4	36.4	56	2	S74259	tiopronin T 3, fast
283	4	36.4	56	2	B72804	gp36 protein - Myc
284	4	36.4	56	2	F81824	hypothetical prote
285	4	36.4	57	2	S62889	non-histone protei
286	4	36.4	59	2	S74260	tiopronin T1f, fast
287	4	36.4	60	2	A30160	H+-transporting AT
288	4	36.4	60	2	D85758	hypothetical prote
289	4	36.4	61	2	C82812	hypothetical prote
290	4	36.4	61	2	F81844	hypothetical prote
291	4	36.4	61	2	T19658	hypothetical prote
292	4	36.4	62	2	T30977	hypothetical prote
293	4	36.4	62	2	T10302	hypothetical prote
294	4	36.4	64	2	F83334	hypothetical prote
295	4	36.4	64	2	D82032	hypothetical prote
296	4	36.4	64	2	E81008	hypothetical prote
297	4	36.4	66	2	H81813	hypothetical prote
298	4	36.4	67	2	S48660	tiopronin T 1, fast
299	4	36.4	67	2	A85736	hypothetical prote
300	4	36.4	67	2	B55855	tall protein - pha
301	4	36.4	67	2	D64894	hypothetical prote
302	4	36.4	68	2	T45155	hypothetical prote
303	4	36.4	69	2	C84275	hypothetical prote
304	4	36.4	69	2	S51861	hypothetical prote
305	4	36.4	70	2	S25190	pillin precursor -
306	4	36.4	70	2	T46343	hypothetical prote
307	4	36.4	71	2	A60325	cytochrome-c oxida
308	4	36.4	71	2	E75316	hypothetical prote
309	4	36.4	71	2	D82541	hypothetical prote
310	4	36.4	72	2	T17585	hypothetical prote
311	4	36.4	73	2	T68536	pepsinogen - huma
312	4	36.4	73	2	A36260	alpha-2-macroglobu
313	4	36.4	73	2	B86279	hypothetical prote
314	4	36.4	74	2	S42233	flxg protein - Azo
315	4	36.4	74	2	E69871	hypothetical prote
316	4	36.4	74	2	H70070	hypothetical prote
317	4	36.4	75	2	B64988	hypothetical 8.3 k
318	4	36.4	75	2	A85858	hypothetical prote
319	4	36.4	76	2	H70576	hypothetical prote
320	4	36.4	76	2	T03810	hypothetical prote
321	4	36.4	77	2	C86842	30S ribosomal prot
322	4	36.4	77	2	B75454	hypothetical prote
323	4	36.4	79	2	T44149	hypothetical prote
324	4	36.4	79	2	T26118	hypothetical prote
325	4	36.4	81	2	D69082	ribosomal protein
326	4	36.4	81	2	S28041	hemp protein - yer
327	4	36.4	83	2	S66246	nuclear factor int
328	4	36.4	83	2	A46348	hypothetical prote
329	4	36.4	83	2	A33854	outer membrane lip
330	4	36.4	84	1	MMBPMB	gene P protein - p
331	4	36.4	84	2	T27174	hypothetical prote
332	4	36.4	85	2	IC4165	hypothetical prote
333	4	36.4	86	1	OC807	cytochrome-c oxida
334	4	36.4	86	2	T07204	hypothetical prote
335	4	36.4	87	1	DA5345	trans-activating t
336	4	36.4	87	2	S44877	zC21.5 protein - C
337	4	36.4	88	2	T36450	hypothetical prote
338	4	36.4	89	2	S15467	hypothetical prote
339	4	36.4	89	2	E84167	hypothetical prote
340	4	36.4	89	2	D72581	hypothetical prote
341	4	36.4	90	1	OOBECF	legume protein -
342	4	36.4	90	2	C85516	probable activator
343	4	36.4	90	2	T33692	hypothetical prote
344	4	36.4	91	1	W4WL58	E4 protein - human
345	4	36.4	91	2	D64066	host factor I - Ha
346	4	36.4	91	2	D86503	hypothetical prote
347	4	36.4	91	2	T30512	hypothetical prote
348	4	36.4	91	2	H75623	hypothetical prote
349	4	36.4	91	2	C72119	hypothetical prote
350	4	36.4	93	2	T06803	probable dehydrin
351	4	36.4	94	2	T06417	pathogenesis-relat
352	4	36.4	94	2	JO1164	Tat protein - Maed
353	4	36.4	94	2	S67047	SM21 protein - yea
354	4	36.4	95	2	SA9571	Ig lambda chain pr
355	4	36.4	95	2	G30603	T-cell receptor al
356	4	36.4	95	2	UC4760	SM73 protein - hum
357	4	36.4	95	2	T45315	hypothetical prote
358	4	36.4	96	2	B41898	dextranucrase (EC
359	4	36.4	97	2	H96743	probable Dnaa prot
360	4	36.4	97	2	S30990	gene 45 protein -
361	4	36.4	98	1	OC806A	cytochrome-c oxida
362	4	36.4	98	2	S36068	Ig lambda chain -
363	4	36.4	98	2	S36047	Ig lambda chain -
364	4	36.4	98	2	S36048	Ig lambda chain -
365	4	36.4	98	2	S36050	Ig lambda chain -
366	4	36.4	98	2	SA9414	major carboxysome
367	4	36.4	98	2	S05527	capsaicin - Phyto
368	4	36.4	98	2	S06671	cinanomoin - Phyto
369	4	36.4	99	2	S36051	Ig lambda chain -
370	4	36.4	99	2	S36052	Ig lambda chain -
371	4	36.4	99	2	S36053	Ig lambda chain -
372	4	36.4	99	2	S36054	Ig lambda chain -
373	4	36.4	99	2	S36058	Ig lambda chain -
374	4	36.4	99	2	S03525	T-cell receptor al
375	4	36.4	99	2	S42360	capsaicin - Phyto
376	4	36.4	99	2	H86580	CP656 hypothetical
377	4	36.4	99	2	E72043	conserved hypothet
378	4	36.4	100	1	MMWGN3	11k protein - narc
379	4	36.4	101	2	A36869	probable export pr
380	4	36.4	101	2	C69930	hypothetical prote
381	4	36.4	101	2	S30970	gene 25 protein -
382	4	36.4	101	2	G72802	gp25 protein - Myc
383	4	36.4	101	2	F85812	flagellar biosynthe
384	4	36.4	102	2	B48360	orf1 upstream of m
385	4	36.4	103	1	LIH0NM	Ig lambda chain V-
386	4	36.4	103	2	A38923	Al type amyloid fl
387	4	36.4	103	2	S62052	hypothetical prote
388	4	36.4	103	2	A71197	hypothetical prote
389	4	36.4	103	2	T40081	very hypothetical
390	4	36.4	103	2	T49368	hypothetical prote
391	4	36.4	103	2	D82165	hypothetical prote
392	4	36.4	103	2	G81399	2-oxoglutarate syn
393	4	36.4	104	1	ASLJCV	vpr protein - huma
394	4	36.4	104	1	ASLJCV	vpr protein - huma

103	5	45.5	343	2	A45073	Gal beta 1,3galNAc	176	5	45.5	610	2	T17775	protein kinase hom
104	5	45.5	344	1	S35140	probable ketol-act	177	5	45.5	613	2	S48557	hypothetical prote
105	5	45.5	352	2	T36719	probable integral	178	5	45.5	613	2	T42671	hypothetical prote
106	5	45.5	354	2	S65687	(A+T)-stretch-bind	179	5	45.5	624	2	JC5471	regulatory protein
107	5	45.5	357	2	F70898	probable gfp-bind1	180	5	45.5	638	1	TOBCDK	dnak-type molecu
108	5	45.5	363	2	I64150	probable gfp-bind1	181	5	45.5	638	2	F85481	hypothetical prote
109	5	45.5	373	2	T451747	GDP-D-mannose-4,6-	182	5	45.5	638	2	C64156	hypothetical prote
110	5	45.5	373	2	E82570	chorismate mutase	183	5	45.5	650	2	T44660	nitrous-oxide redu
111	5	45.5	375	2	F75467	probable lipoprote	184	5	45.5	650	2	T36419	hypothetical prote
112	5	45.5	378	2	A47300	cell adhesion prot	185	5	45.5	656	1	B44954	hypothetical prote
113	5	45.5	380	2	S51797	vasodilator-stimul	186	5	45.5	656	2	B86033	hypothetical prote
114	5	45.5	382	2	T14336	RAD23 protein, iso	187	5	45.5	659	2	E86313	hypothetical prote
115	5	45.5	385	2	T29315	hypothetical prote	188	5	45.5	667	2	A48579	hypothetical prote
116	5	45.5	388	2	T26372	hypothetical prote	189	5	45.5	681	2	S36534	trypothozite surfac
117	5	45.5	398	2	S50507	excision repair pr	190	5	45.5	701	2	S72722	protein secretion
118	5	45.5	408	2	C70379	hypothetical prote	191	5	45.5	714	2	G71963	hypothetical prote
119	5	45.5	412	2	A48702	2-methyl-branched-	192	5	45.5	714	2	H64543	hypothetical prote
120	5	45.5	412	2	C72548	probable dihydrol	193	5	45.5	718	2	A81122	hypothetical prote
121	5	45.5	412	2	S17758	tubulin beta chain	194	5	45.5	728	2	T26607	hypothetical prote
122	5	45.5	414	1	JN0826	granulatin polyket	195	5	45.5	733	2	G84668	hypothetical prote
123	5	45.5	415	1	S05394	hypothetical prote	196	5	45.5	768	2	F83060	hypothetical prote
124	5	45.5	416	2	T34479	histidyl-tRNA synt	197	5	45.5	770	2	G02228	DOC-2 - human
125	5	45.5	431	2	G75017	hypothetical prote	198	5	45.5	787	2	A44452	translation initia
126	5	45.5	445	2	F83881	hypothetical prote	199	5	45.5	800	2	S37387	internalin A precu
127	5	45.5	445	2	E70360	glucose inhibited	200	5	45.5	820	2	T51510	hypothetical prote
128	5	45.5	447	2	S37048	cysteine proteinas	201	5	45.5	824	2	T23923	hypothetical prote
129	5	45.5	448	2	T27955	hypothetical prote	202	5	45.5	837	2	A57542	p96 protein - mous
130	5	45.5	448	2	S39348	26S ATP/ubiquitin-	203	5	45.5	863	2	T38016	hypothetical prote
131	5	45.5	448	2	T01570	hypothetical prote	204	5	45.5	876	2	S20531	importin beta-1 su
132	5	45.5	451	2	D47691	UDP-N-acetylmuramo	205	5	45.5	891	2	H83218	hypothetical prote
133	5	45.5	451	2	A81206	sugar transporter,	206	5	45.5	894	2	D85569	heme acquisition p
134	5	45.5	451	2	E81781	probable integral	207	5	45.5	894	2	B42372	histidine kinase k
135	5	45.5	456	2	T46986	hypothetical prote	208	5	45.5	927	2	T24031	hypothetical prote
136	5	45.5	457	2	T24962	odr-7 protein - Ca	209	5	45.5	941	2	A86404	probable protein A
137	5	45.5	459	1	KIXEFV	major capsid prote	210	5	45.5	943	2	JC4081	sucrase/fructanase
138	5	45.5	470	2	S50854	DNA-1 protein - hu	211	5	45.5	971	2	T33907	hypothetical prote
139	5	45.5	471	2	T48743	probable 26S ATP/u	212	5	45.5	986	2	B81675	hypothetical prote
140	5	45.5	476	2	B44997	mezozoite surface	213	5	45.5	1012	2	I53172	polymorphic membra
141	5	45.5	479	2	A38578	protein kinase 2 (	214	5	45.5	1027	2	A56533	RAE-28 - mouse
142	5	45.5	479	2	G83259	hypothetical prote	215	5	45.5	1040	1	T46931	alpha-mannosidase
143	5	45.5	480	2	G96723	hypothetical prote	216	5	45.5	1040	1	A38306	alpha-mannosidase
144	5	45.5	485	2	F83593	outer membrane pro	217	5	45.5	1041	2	A81960	probable DNA helic
145	5	45.5	489	2	T04596	hypothetical prote	218	5	45.5	1049	2	T30525	alpha-mannosidase
146	5	45.5	500	2	T19525	hypothetical prote	219	5	45.5	1051	1	JW0051	serine/chreonine-s
147	5	45.5	504	2	T13475	hypothetical prote	220	5	45.5	1061	1	S27311	ribonuclease E (EC
148	5	45.5	508	1	A30007	dolichyl-diphospho	221	5	45.5	1061	2	B85671	hypothetical prote
149	5	45.5	508	1	ISPRSS	protein disulfide-	222	5	45.5	1070	2	S19686	alpha-glucosidase
150	5	45.5	508	1	ISRUS	protein disulfide-	223	5	45.5	1098	2	B70232	hypothetical prote
151	5	45.5	509	1	ISMSSS	protein disulfide-	224	5	45.5	1151	2	T18535	high molecular mas
152	5	45.5	510	1	A38362	protein disulfide-	225	5	45.5	1151	2	T30936	reverse transcript
153	5	45.5	510	1	ISBOSS	protein disulfide-	226	5	45.5	1161	2	G81915	hypothetical prote
154	5	45.5	510	2	A54798	Vg-1-related prote	227	5	45.5	1177	2	T16594	hypothetical prote
155	5	45.5	513	2	F83162	nitrate reductase	228	5	45.5	1218	2	J50069	hypothetical pi op
156	5	45.5	515	1	ISCHSS	protein disulfide-	229	5	45.5	1224	2	T14007	microtubule-associ
157	5	45.5	520	2	T04591	ferulate-5-hydroxy	230	5	45.5	1237	2	A54080	protein-tyrosine-p
158	5	45.5	527	2	T07607	phosphate transpor	231	5	45.5	1301	1	A41622	protein-tyrosine-p
159	5	45.5	528	2	T05724	probable inorganic	232	5	45.5	1329	2	A84525	probable retroelem
160	5	45.5	533	2	S69336	arylsulfatase (EC	233	5	45.5	1356	2	T16718	hypothetical prote
161	5	45.5	536	2	D84325	HtrII transducer (	234	5	45.5	1403	2	S24548	homeotic protein p
162	5	45.5	540	2	T00646	hypothetical prote	235	5	45.5	1409	2	T37188	presynaptic activi
163	5	45.5	545	2	PM0005	chaperonine 60K al	236	5	45.5	1416	2	T20823	hypothetical prote
164	5	45.5	545	2	I48951	oncotelet antigen	237	5	45.5	1520	2	B82274	conserved hypotet
165	5	45.5	553	1	H70786	probable dihydrol	238	5	45.5	1612	2	S59969	DNA topoisomerase
166	5	45.5	554	2	S46775	hypothetical prote	239	5	45.5	1621	2	A82255	hypothetical prote
167	5	45.5	576	2	G69880	hypothetical prote	240	5	45.5	1640	2	D86798	prophage p13 prote
168	5	45.5	576	2	T46385	hypothetical prote	241	5	45.5	1785	2	A45546	major merozoit su
169	5	45.5	578	2	T40984	transcription fact	242	5	45.5	1785	2	T21558	hypothetical prote
170	5	45.5	581	2	E75383	conserved hypotet	243	5	45.5	1825	2	T52521	related to SEN1 pr
171	5	45.5	583	2	S10712	acetylcholinestera	244	5	45.5	1870	2	S37671	MHC class III hist
172	5	45.5	583	2	T16007	hypothetical prote	245	5	45.5	1872	2	S36152	MHC class III hist
173	5	45.5	584	2	S55106	probable membrane	246	5	45.5	1876	2	T13801	phosphoinositide 3
174	5	45.5	587	1	A36936	nitrogenase (EC 1.	247	5	45.5	1899	2	T32732	PAM C-terminal int
175	5	45.5	597	2	S40998	hypothetical prote	248	5	45.5	1919	2	T42098	PAM interacting pr

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:34:31 ; Search time 33.87 Seconds  
(without alignments)  
24.739 Million cell updates/sec

Title: US-09-372-036-29  
Perfect score: 11  
Sequence: 1 OOOQAPKAPTE 11

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size: 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	100.0	484	2 A41487	protein p60 precu
2	6	54.5	343	2 T29819	hypothetical prote
3	6	54.5	371	2 T22894	hypothetical prote
4	6	54.5	606	2 T37217	probable ATP/GTP b
5	6	54.5	663	2 G81384	succinate dehydrog
6	6	54.5	861	2 H83012	periplasmic glucan
7	6	54.5	890	2 A48753	NPAT transcription
8	6	54.5	995	2 T22942	hypothetical prote
9	6	54.5	1001	2 S74544	translation initia
10	6	54.5	1052	2 T14343	zinc finger RNA bi
11	6	54.5	1323	1 RRMGSW	RNA-directed RNA p
12	6	54.5	2022	2 T43214	ovrl protein - nem
13	6	54.5	2121	2 A59233	myosin VII-like pr
14	5	45.5	46	2 S07073	arabinogalactan pr
15	5	45.5	67	2 I47395	histone H1 I-1 (cl
16	5	45.5	70	2 A47395	histone H1 I-1 (N-
17	5	45.5	83	2 A44465	sodium ion pump ox
18	5	45.5	95	2 S36634	ptl2 protein - Ant
19	5	45.5	95	2 G75374	hit family protein
20	5	45.5	118	2 B49905	protein secretion
21	5	45.5	119	2 B81363	Hit-family protein
22	5	45.5	123	2 S49484	histone H2B - sea
23	5	45.5	127	2 S09877	hypothetical prote
24	5	45.5	127	2 A72488	hypothetical prote
25	5	45.5	128	2 G64155	Yhc protein - Hae
26	5	45.5	129	1 CCDV3	cytochrome c3 prec
27	5	45.5	130	1 CCDV3M	large-conductance
28	5	45.5	136	2 H82438	dehydrin 7 - barle
29	5	45.5	137	2 S35752	

30	5	45.5	139	2 S05544	dehydrin 8 - barle
31	5	45.5	139	2 T05715	dehydrin - barley
32	5	45.5	148	2 C81214	comE-related prot
33	5	45.5	157	2 F83464	flagellar motor sw
34	5	45.5	159	2 F81813	hypothetical prote
35	5	45.5	166	2 I39546	hypothetical prote
36	5	45.5	174	2 A71208	hypothetical prote
37	5	45.5	175	2 A26882	ptl2 hypothetical
38	5	45.5	176	2 T48699	hypothetical prote
39	5	45.5	181	2 F70588	hypothetical prote
40	5	45.5	182	2 T11906	NADH dehydrogenase
41	5	45.5	183	2 E75435	pyrimidine operon
42	5	45.5	191	2 T42232	histone H1.2 - Cae
43	5	45.5	193	2 S73713	Mg319 homolog H08
44	5	45.5	202	1 S17244	adenylylsulfate ki
45	5	45.5	205	2 H81057	comE operon protei
46	5	45.5	207	2 T49218	hypothetical prote
47	5	45.5	207	2 D83944	SOS regulon transc
48	5	45.5	212	2 H75317	translational initia
49	5	45.5	214	2 S70734	adenylylate kinase (
50	5	45.5	216	2 T38907	hypothetical prote
51	5	45.5	218	2 H82758	hypothetical prote
52	5	45.5	223	2 T16654	hypothetical prote
53	5	45.5	224	2 T43331	clathrin light cha
54	5	45.5	225	2 T15228	probable histone B
55	5	45.5	229	2 T40789	clathrin light cha
56	5	45.5	231	2 C64990	ferredoxin-type pr
57	5	45.5	231	2 H85859	ferredoxin-type pr
58	5	45.5	236	2 S22322	histone H1 - wheat
59	5	45.5	239	2 S28630	opacity protein op
60	5	45.5	241	2 F42600	probable octopline
61	5	45.5	244	2 S40436	histone H1 - midge
62	5	45.5	252	1 F70647	NADH dehydrogenase
63	5	45.5	261	2 JN0747	histone H1-I - Vol
64	5	45.5	262	2 T31468	hypothetical prote
65	5	45.5	266	2 S16616	opacity protein op
66	5	45.5	269	2 D82060	conserved hypothet
67	5	45.5	270	2 S75051	lactose transport
68	5	45.5	270	2 B83927	hypothetical prote
69	5	45.5	279	1 SUM71Y	thermolase (EC 3.4
70	5	45.5	280	2 T09576	Lbx1 transcription
71	5	45.5	283	2 S25769	lipH protein - pse
72	5	45.5	288	2 C83289	lipase modulator p
73	5	45.5	289	2 T17957	hypothetical prote
74	5	45.5	292	2 B42822	cyclin D3 - human
75	5	45.5	296	2 A56943	sensory/motor neur
76	5	45.5	296	2 T20005	hypothetical prote
77	5	45.5	298	1 C37841	probable site-spec
78	5	45.5	299	2 G86649	drp-E-ribose sy
79	5	45.5	301	2 C85049	hypothetical prote
80	5	45.5	303	2 C75289	hypothetical prote
81	5	45.5	304	2 A61128	probable lysR-famI
82	5	45.5	307	2 F84502	En/Spm transposon
83	5	45.5	313	2 T29525	desiccation-relate
84	5	45.5	315	2 T29525	hypothetical prote
85	5	45.5	320	2 S78604	taurin-binding pro
86	5	45.5	320	2 G85531	taurine transport
87	5	45.5	320	2 S61586	probable membrane
88	5	45.5	321	2 D70833	hypothetical prote
89	5	45.5	321	2 C71463	probable p-loop AT
90	5	45.5	323	2 I39867	microbial serine p
91	5	45.5	327	2 D83674	phenylacetic acid
92	5	45.5	330	2 T25169	hypothetical prote
93	5	45.5	330	2 T42947	hypothetical prote
94	5	45.5	331	2 B85830	spherulin 4 precu
95	5	45.5	332	2 S03871	hypothetical prote
96	5	45.5	333	2 A85620	probable nitrate t
97	5	45.5	333	2 G64833	kecol-acid reducto
98	5	45.5	338	2 B83059	kecol-acid reducto
99	5	45.5	339	2 S24161	kecol-acid reducto
100	5	45.5	340	2 B86778	kecol-acid reducto
101	5	45.5	340	2 G81411	p3 protein - grape
102	5	45.5	341	1 P3V6CF	



**This Page Blank (uspto)**

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TAPKA 8  
 |||||  
 Db 92 tapka 96

RESULT 50  
 AAR54838  
 ID AAR54838 standard; Protein: 279 AA.  
 XX  
 AC AAR54838;  
 XX  
 DT 11-JAN-1995 (first entry)  
 XX  
 DE Thermilase (P229X).  
 XX  
 KW Subtilisin; mutant; variant; enzyme; protease; substrate;  
 KM thermilase.  
 XX  
 OS Thermoterranomyces valgaris.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 229  
 FT /label= Ala, Leu, Met, Gin, Val, Ser  
 FT /note= "pref. Ala"  
 XX  
 PN US5316941-A.  
 XX  
 PD 31-MAY-1994.  
 XX  
 PF 06-JAN-1989; 89US-0294340.  
 XX  
 PR 06-JAN-1988; 89US-0294340.  
 PR 29-APR-1992; 92US-0876818.  
 XX  
 PA (GENV ) GENENCOR INC.  
 XX  
 PI Bott RR, Caldwell RM, Estell DA, Graycar TP;  
 XX  
 DR WPI; 1994-176274/21.  
 XX  
 PT New DNA encoding subtilisin mutated at position 225 - producing  
 PT protease with altered substrate specificity and reactivity  
 XX  
 PS Disclosure: Fig 3b; 13pp: English.  
 XX  
 CC DNA encoding subtilisin in which the amino acid equiv. to Pro(225)  
 CC of B. amyloliquefaciens subtilisin has been replaced by Ala, Leu,  
 CC Met, Gin, Val or Ser (pref. Ala) is claimed. Compared with wild-  
 CC type subtilisin, the variants have different catalytic and/or  
 CC Michaelis constants, i.e. different substrate reactivity and/or  
 CC specificity, partic. at least a 2-fold increase or decrease in the  
 CC kcat/Km ratio. Changing Pro to Ala gives a mutant enzyme with  
 CC greater activity for ester as opposed to anilide substrates.  
 XX  
 SQ Sequence 279 AA;

Query Match 45.5%; Score 5; DB 15; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TAPKA 8  
 |||||  
 Db 92 tapka 96

Search completed: August 15, 2001, 12:35:42  
 Job time: 167 sec

PT Isolated from culture medium of fungus *Tritirachium album* or  
 PT produced by recombinant DNA techniques  
 XX  
 PS ; Fig 9a-9c; 74pp; English.  
 CC Fig. 9 represents a comparison of the AA sequences of protease TW3 with  
 CC those of proteinase K, subtilisin novo, subtilisin Carlsberg, subtilisin  
 CC DY and thermolase. The AA sequence of the mature TW3 protein  
 CC as determined from the nucleotide sequence has approximately 90%  
 CC homology with that of proteinase K. There are certain positions where  
 CC the TW3 AA sequence resembles that in subtilisins, but not to proteinase  
 CC K. For example, at positions 143, a methionine residue occurs in all  
 CC subtilisins as well as in protease TW3, while a leucine residue is  
 CC present at that position in proteinase K. Similarly at position 219, an  
 CC alanine residue is present in protease TW3 and subtilisins, but not in  
 CC proteinase K. In addition, the AA fragment, Ser-Thr-, is absent from  
 CC AAP80847-49 and AAP80851 at position 226 and 227. Purified serine protease (SP),  
 CC can be used in detergents and cleansers or spot removers, as a depilatory  
 CC in tanning and also in the food industry for the prepn. of protein  
 CC hydrolysates and in serology for the detection of incomplete antibodies.  
 XX  
 SQ Sequence 278 AA;  
 Query Match 45.5%; Score 5; DB 9; Length 278;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 4 TAPKA 8  
 |||||  
 Db 92 tapka 96  
 RESULT 48  
 AAR90544 standard; protein: 278 AA.  
 ID AAR90544  
 XX  
 AC AAR90544;  
 XX  
 DT 08-AUG-1996 (first entry)  
 XX  
 DE pUG4-5-CDK-BP clone #125 derived CDK4 binding protein.  
 XX  
 KM Cell cycle; CDK4; regulation; G1 phase; proliferation; tumorigenesis;  
 KM cyclin dependent kinase; differentiation; CDK4 inhibitor; agonist;  
 KM antagonist.  
 XX  
 OS Synthetic.  
 OS  
 XX  
 PN MO9533819-A2.  
 XX  
 PD 14-DEC-1995.  
 XX  
 PF 02-JUN-1995; 95WO-US07113.  
 XX  
 PR 02-JUN-1994; 94US-0253155.  
 XX  
 PA (MITO-) MITOTIX INC.  
 XX  
 PI Draetta G, Gyuris J;  
 XX  
 DR WPI; 1996-040227/04.  
 DR N-PSDB; AAT12173.  
 XX  
 XX Cyclin-dependent kinase-4 binding protein - used in the isolation of  
 PT (ant)agonists of cell cycle regulation.  
 XX  
 PS Claim 1; Page 81-82; 115pp; English.  
 XX  
 CC AAR90533-R90556 are cyclin dependent kinase 4 (CDK4) binding proteins.  
 CC encoded by clones of the plasmid pUG4-5-CDKBP. CDK4 binding proteins  
 CC (CDK4-BP) may be used in an assay for screening test compounds as

CC Inhibitors of CDK/CDK4-BP interaction. The complexes formed by CDK4  
 CC and D-type cyclins are strongly implicated in the control of the early  
 CC G1 phase of the cell cycle and are strong candidates for controlling  
 CC and/or preventing tumorigenesis and the onset of cancer. Nucleic acids  
 CC encoding CDK4-BP or fragments of these may be used as probes/primers  
 CC to diagnose the presence or absence of genetic lesions in a gene  
 CC encoding 1 of the 24 CDK4-BP, and hence to diagnose the risk for a  
 CC subject of developing a cell-proliferation associated disorder (e.g.  
 CC cancer).  
 XX  
 SQ Sequence 278 AA;  
 Query Match 45.5%; Score 5; DB 17; Length 278;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 4 TAPKA 8  
 |||||  
 Db 57 tapka 61  
 RESULT 49  
 AAP80274  
 ID AAP80274 standard; protein: 279 AA.  
 XX  
 AC AAP80274;  
 XX  
 DT 10-SEP-1990 (first entry)  
 XX  
 DE Amino acid sequence of *Bacillus thermolase* subtilisin.  
 XX  
 KM Enzyme mutant; catalytic function; *Bacillus thermolase* subtilisin.  
 XX  
 OS *Bacillus thermolase*.  
 OS  
 XX  
 PN WO8807578-A.  
 XX  
 PD 06-OCT-1988.  
 XX  
 PF 30-MAR-1988; 88WO-US01078.  
 XX  
 PR 02-APR-1987; 87US-0034085.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Wells JA, Carter PJ;  
 XX  
 DR WPI; 1988-292862/41.  
 XX  
 PT Enzyme mutants with modified catalytic functional gp. -  
 PT react with modified substrates which contain moiety to  
 PT complement catalytic functionality  
 XX  
 PS Disclosure; Fig 3b; 90pp; English.  
 XX  
 CC The patent is for a novel enzyme mutant (EM), not found in nature,  
 CC derived by the replacement or modification, in a precursor enzyme, of at  
 CC least one catalytic gp. of an amino acid residue which when in contact  
 CC with a selected region of a polypeptide substrate, functions  
 CC catalytically. The precursor enzyme may be, eg oxido-reductases,  
 CC transferases, hydrolases, lyases, isomerases or ligases. EM have a  
 CC catalytic preference for substrates which provide the replaced or  
 CC modified functional gp. or its equivalent such that the substrate  
 CC together with the enzyme mutant assists in its own catalysis.  
 CC In Fig 3b B. amyloliquefaciens subtilisin and thermolase are aligned to  
 CC compare conserved residues.  
 XX  
 SQ Sequence 279 AA;  
 Query Match 45.5%; Score 5; DB 9; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;

```

FT      /note= "unidentified amino acid"
PR      Misc-difference 16
XX      /note= "unidentified amino acid"
XX      W09709994-A1.
XX      20-MAR-1997.
XX      16-SEP-1996; 96WO-US14819.
XX      15-SEP-1995; 95US-0529055.
XX      (UABR-) UAB RES FOUND.
XX      Briles DE, Brooks-Walter A, Crain MJ, Hollingshead S;
PI      McDaniel LS, Swiatlo E, Tart R, Yother J;
XX      WPI: 1997-202002/18.
XX      Streptococcus pneumoniae surface protein PspC and truncated PspA -
PT      used in vaccines for protecting animals against S.pneumoniae
PT      infection
XX      Example 6; Fig 13; 296pp; English.
XX      This sequence shows the central portion, including the C-terminus
CC      of the alpha-helix region and some of the proline-rich region, of
CC      pneumococcal surface protein A (PspA) of Streptococcus pneumoniae
CC      strain ATCC6303. Comparison of the N-terminal and central regions
CC      (AAW14333-57 and AAW14562-91) of PspA polypeptides from different
CC      pneumococcal strains can be used to divide the strains into several
CC      families based on sequence homologies. PspA polypeptides, or
CC      fragments of them, can be used in vaccines to protect animals
CC      against S. pneumoniae infection and hence for the prevention of
CC      diseases such as otitis media, meningitis, bacteraemia and pneumonia.
CC      The sequence of the 3' half of the PspA alpha-helical region and the
CC      immediate 5' tip of the coding sequence are likely to be the critical
CC      sequences for predicting PspA cross-reactions and vaccine
CC      composition.
XX      Sequence 215 AA:
SQ

```

Query Match 45.5%; Score 5; DB 18; Length 215;  
Best Local Similarity 100.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY      4 TAPKA 8
DB      125 tapka 129

```

RESULT 46  
AAW80691  
ID AAW80691 standard; Protein: 243 AA.  
XX  
AC AAW80691;  
XX  
DT 24-DEC-1998 (first entry)  
XX  
DE S. pneumoniae protein of unknown function.  
XX  
KM Streptococcus pneumoniae protein; recombinant; gene expression; DNA chip;  
XX virulence; antibody; infection; detection; treatment; hypothetical;  
XX cell wall biosynthetic, external target; minimal gene set protein.  
XX  
OS Streptococcus pneumoniae.  
XX  
XX W09826072-A1.  
XX  
XX 18-JUN-1998.  
XX  
XX 09-DEC-1997; 97WO-US22578.

```

XX      13-DEC-1996; 96US-0036281.
XX      (ELL) LILLY & CO ELI.
XX      Balz RH, Burgett SG, Dehoff BS, Hoskins JA, Jaskunas SR;
XX      PI Mills BJ, Norris FH, Peery RB, Rockey PK, Rostock PR;
XX      PI Skatrud PL, Smith MC, Solenberg PJ, Treadway PJ;
XX      PI Young Bellido ML;
XX      WPI: 1998-348529/30.
XX      N-PSDB: AAV65255.
XX      Streptococcus pneumoniae nucleic acid sequences - used in DNA chips
PT      for evaluating gene expression, and identification of virulence
PT      genes
XX      Claim 3; Pages 279-280; 333pp; English.
XX      This sequence represents a Streptococcus pneumoniae protein of unknown
CC      function. The invention provides DNA sequences (AAV65201 to AAV65304)
CC      from the Streptococcus pneumoniae genome and corresponding protein
CC      sequences (AAW80605 to AAW80728). The protein sequences are classified as
CC      hypothetical, cell wall biosynthetic, external target, or minimal gene
CC      set proteins. A recombinant host containing a vector comprising any of
CC      the above nucleic acids can be used for the recombinant expression of the
CC      proteins. The invention also provides a DNA chip having arrayed on it at
CC      least 15 base pair fragment of any one or more of these DNA sequences.
CC      The DNA chip can be used methods for evaluating gene expression in S.
CC      pneumoniae and for identifying virulence genes in S. pneumoniae.
CC      Antibodies that selectively bind to the above proteins or peptide
CC      fragments can be used to treat S. pneumoniae infection. The antibodies
CC      can also be used to detect S. pneumoniae cells.
XX      Sequence 243 AA:
SQ

```

Query Match 45.5%; Score 5; DB 19; Length 243;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY      7 KAPTE 11
DB      120 kapte 124

```

RESULT 47  
AAP80850  
ID AAP80850 standard; protein: 278 AA.  
XX  
AC AAP80850;  
XX  
DT 10-SEP-1990 (first entry)  
XX  
DE sequence of subtilisin thermilase.  
XX  
KM Serine protease TW3; proteinase K; subtilisin thermilase.  
XX  
XX W08807581-A.  
XX  
XX 06-OCT-1988.  
XX  
XX 28-MAR-1988; 88WO-US01040.  
XX  
XX 03-MAR-1987; 87US-0035816.  
XX  
XX (AMGE ) AMGEN INC.  
XX  
XX Samal BB, Stabinsky Y;  
XX  
XX WPI: 1988-292865/41.  
XX  
XX New purified serine protease -

XX PN WO200006736-A2.  
 XX XX 10-FEB-2000.  
 PD XX  
 PF 27-JUL-1999; 99WO-GB02444.  
 XX XX  
 XX 27-JUL-1998; 98GB-0016335.  
 PR 19-MAR-1999; 99US-0125163.  
 XX XX  
 PA (MICR-) MICROBIAL TECHNIQS LTD.  
 XX PI  
 XX PI Le Page RWF, Wells JM, Hanniffy SB;  
 DR WPI: 2000-195299/17.  
 XX  
 PT New Group B Streptococcus protein, useful as vaccine, for diagnosis of  
 PT Streptococcal infections and for screening of antibodies or affibodies  
 XX  
 PS Claim 1; Fig 1: 123pp; English.  
 XX  
 CC AAA05803 to AAA05872 encode proteins, polypeptides and peptides (given  
 CC in AA91375 to AA91343) isolated from Group B Streptococcus (GBS), also  
 CC known as Streptococcus agalactiae. The GBS polynucleotides and  
 CC polypeptides have antibacterial activity. Immunogenic compositions  
 CC comprising GBS polynucleotides or polypeptides can be used as vaccines  
 CC and for the treatment or prophylaxis of GBS infection. The  
 CC polynucleotides and polypeptides can also be used in the detection of GBS  
 CC and for screening DNA encoding bacterial cell envelope associated or  
 CC secreted antigens in gram positive bacteria. AAA05873 to AAA05941  
 CC represent primers used in the exemplification of the present invention.  
 XX  
 SQ Sequence 206 AA;  
 Query Match 45.5%; Score 5; DB 21; Length 206;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 6 PKAPT 10  
 Db 144 pkapt 148  
 DE  
 RESULT 44  
 AAB42897  
 ID AAB42897 standard; Protein: 209 AA.  
 AC AAB42897;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF2661 polypeptide sequence SEQ ID NO:5332.  
 XX  
 KW Human: open reading frame; ORFX; detection; cytosolic; hepatotropic;  
 KW vulnery; antipsoxic; antiparkinsonian; noctropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;  
 thrombosis; contraceptive.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX

PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US08621.  
 XX  
 PR 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX PI  
 XX PI Shimkets RA, Leach M;  
 DR WPI: 2000-602362/57.  
 DR N-PSDB: AAC77106.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 PS Claim 11; Page 4497-4498; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytosolic; hepatotropic; vulnery;  
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
 CC antithyroid; and antianemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy.  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 XX  
 SQ Sequence 209 AA;  
 Query Match 45.5%; Score 5; DB 21; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 7 KAPTE 11  
 Db 195 kapte 199  
 DE  
 RESULT 45  
 AAAM14563  
 ID AAAM14563 standard; Protein: 215 AA.  
 AC AAAM14563;  
 XX  
 DT 28-OCT-1997 (first entry)  
 XX  
 DE Streptococcus pneumoniae PspA central region.  
 XX  
 KW PspA; pneumococcal surface protein; vaccine; otitis media;  
 KW meningitis; bacteraemia; pneumonia.  
 XX  
 OS Streptococcus pneumoniae strain ATCC 6303.  
 XX  
 PN Key Location/Qualifiers  
 FT Misc-difference 4

CC chemicals, for modulating fine chemical production in *C. glutamicum* or  
 CC related bacteria (e.g. *Brevibacterium lactofermentum*), the typing or  
 CC identification of *C. glutamicum* or related bacteria, as reference points  
 CC for mapping *C. glutamicum* genome, and as markers for transformation.  
 CC AAF68082 and AAF68082 represent sequencing primers which are used in an  
 CC example from the present invention.

XX Sequence 189 AA;

Query Match 45.5%; Score 5; DB 22; Length 189;

Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
 |||||  
 Db 20 tapka 24

RESULT 41

AA82664

ID AAR82664 standard; Protein; 190 AA.

XX AAR82664;

DT 27-JAN-1996 (first entry)

DE E. coli RNase E.

KW RNase E; antiviral; transgenic plant.

OS Escherichia coli.

PN W09522245-A.

PD 24-AUG-1995.

PF 16-FEB-1995; 95WO-US02058.

PR 18-FEB-1994; 94US-0198973.

PA (CLEV-) CLEVELAND CLINIC FOUND.

PI Sengupta DN, Silverman RH;

DR WPI; 1995-302493/39.

PT New transgenic plants resistant to viral infection contg. 2-5A-dependent

PT RNase - useful in developing products useful in gene therapy against

PT viral disease and cancer.

PS Disclosure; Page 110; 196pp; English.

XX The novel 2-5A dependent RNase of the invention has a mol. wt. of

CC about 84 KDa. Analysis of the AA sequence has revealed several

CC characteristics unique to the 2-5A-dependent RNases.

CC There is limited homology with RNase E, encoded by the altered

CC mRNA stability (ams)/rne gene of *E. coli*.

XX Sequence 190 AA;

OY 1 00QTA 5  
 |||||  
 Db 41 qgqta 45

RESULT 42

AAW94681

ID AAW94681 standard; Protein; 190 AA.

XX AAW94681;

DT 04-MAY-1999 (first entry)

DE Escherichia coli RNase E.

KW 2-5A-dependent RNase; endoribonuclease; interferon; insect cell;

KW 5'-phosphorylated 2',5'-linked oligoadenylate; recombinant protein.

OS Escherichia coli.

PN US5877019-A.

PD 02-MAR-1999.

PF 21-AUG-1996; 96US-0701005.

PR 22-OCT-1993; 93US-0141304.

PR 08-MAR-1993; 93US-0028086.

PR 21-AUG-1996; 96US-0701005.

PA (CLEV-) CLEVELAND CLINIC FOUND.

PI Hassel BA, Silverman RH, Zhou A;

DR WPI; 1999-189652/16.

PT Insect cells expressing 2-5A-dependent RNase enzymes - useful for

PT producing recombinant enzymes

PS Disclosure; Fig 9A; 62pp; English.

XX The present invention describes an insect cell capable of expressing

CC human or mouse 2-5A-dependent RNase (an endoribonuclease that requires

CC 5'-phosphorylated 2',5'-linked oligoadenylates) at levels at least two

CC orders of magnitude higher than those obtained from mammalian cells,

CC where the insect cell includes a recombinant baculovirus expression

CC vector that contains cDNA encoding a 2-5A-dependent RNase. The insect

CC cell can be used to produce recombinant 2-5A-dependent RNase (in amounts

CC of tens of milligrams), allowing meaningful biophysical studies to be

CC conducted on 2-5A-dependent RNase. The present sequence represents

CC Escherichia coli RNase E which has homology with human and mouse

CC 2-5A-dependent RNase.

XX Sequence 190 AA;

OY 1 00QTA 5  
 |||||  
 Db 41 qgqta 45

RESULT 43

AA91339

ID AAY91339 standard; Protein; 206 AA.

XX AAY91339;

DT 30-MAY-2000 (first entry)

DE Group B Streptococcus protein sequence SEQ ID NO:72.

KW Group B Streptococcus; Streptococcus agalactiae; protein antigen;

KW vaccine; screening; immunogen; detection; diagnosis; infection;

KW antibody; antibody; antibacterial.

OS Streptococcus agalactiae.

Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PKAPT 10  
|1111  
Db 148 pkapt 152

RESULT 39  
AAB63253  
ID AAB63253 standard; Protein: 188 AA.

AC AAB63253;  
XX  
DT 26-MAR-2001 (first entry)  
XX

DE Human breast cancer associated antigen protein sequence SEQ ID NO:615.

KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;  
KW cancer associated antigen; cytostatic; cancer vaccine.

OS Homo sapiens.

PN MO200073801-A2.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000MO-US14749.

PR 28-MAY-1999; 99US-0136526.

PR 10-SEP-1999; 99US-0153454.

PA (LUDW-) LUDWIG INST CANCER RES.

XX  
PI OData Y;

DR WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated  
PT antigen precursors, useful for diagnosing and treating a condition  
PT characterized by expression of an abnormal amount of a protein, e.g.

PT cancer -

XX Example 1; Page 483-484; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014

CC represent nucleotide sequences encoding human breast, gastric and

CC prostate cancer associated antigen precursors (CAAP) respectively.

CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970

CC represent human breast, gastric and prostate CAAP protein sequence

CC respectively. CAAPs have cytostatic activity and can be used in the

CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic

CC acids or anti-CAAP antibodies are useful for diagnosing and treating a

CC condition characterized by expression of an abnormal amount of a protein,

XX e.g. cancer.

XX  
SQ Sequence 188 AA;

OY 4 TAPKA 8  
|1111  
Db 87 tapka 91

RESULT 40  
AAB76629  
ID AAB76629 standard; Protein: 189 AA.  
XX  
AC AAB76629;

XX 11-APR-2001 (first entry)  
DT  
XX

DE Corynebacterium glutamicum MCT protein SEQ ID NO:240.

XX Corynebacterium glutamicum; brevbacterium lactofermentum; MCT;

KW membrane construction and membrane transport protein; petroleum spill;

KW hydrocarbon degradation; gram positive aerobic bacterium; marker;

KW identification; microorganism; fine chemical production; transformation;

XX genome mapping; genetic engineering.

XX  
OS Corynebacterium glutamicum.

XX  
PN MO200100805-A2;

XX  
PD 04-JAN-2001.

PF 23-JUN-2000; 2000MO-IB00926.

PR 25-JUN-1999; 99US-0141031.

PR 08-JUL-1999; 99DE-1031454.

PR 08-JUL-1999; 99DE-1031476.

PR 08-JUL-1999; 99DE-1031563.

PR 09-JUL-1999; 99DE-1032122.

PR 09-JUL-1999; 99DE-1032124.

PR 09-JUL-1999; 99DE-1032125.

PR 09-JUL-1999; 99DE-1032128.

PR 09-JUL-1999; 99DE-1032180.

PR 09-JUL-1999; 99DE-1032182.

PR 09-JUL-1999; 99DE-1032190.

PR 09-JUL-1999; 99DE-1032191.

PR 09-JUL-1999; 99DE-1032209.

PR 09-JUL-1999; 99DE-1032212.

PR 09-JUL-1999; 99DE-1032227.

PR 09-JUL-1999; 99DE-1032228.

PR 09-JUL-1999; 99DE-1032229.

PR 14-JUL-1999; 99DE-1032927.

PR 14-JUL-1999; 99DE-1033005.

PR 14-JUL-1999; 99DE-1033006.

PR 27-AUG-1999; 99DE-1040764.

PR 27-AUG-1999; 99DE-1040765.

PR 27-AUG-1999; 99DE-1040766.

PR 27-AUG-1999; 99DE-1040830.

PR 27-AUG-1999; 99DE-1040831.

PR 27-AUG-1999; 99DE-1040832.

PR 31-AUG-1999; 99DE-1041378.

PR 31-AUG-1999; 99DE-1041379.

PR 31-AUG-1999; 99DE-1041379.

PR 03-SEP-1999; 99DE-1042077.

PR 03-SEP-1999; 99DE-1042078.

PR 03-SEP-1999; 99DE-1042079.

PR 03-SEP-1999; 99DE-1042080.

XX (BADI ) BASF AG.

XX  
PA Pompejus M, Kroegeer B, Schroeder H, Zeider O, Haberhauer G;

XX  
PI WPI; 2001-071486/08.

XX  
DR N-PSDB; AAF67862.

XX  
XX Corynebacterium glutamicum nucleic acids encoding membrane construction

PT and membrane transport proteins or their portions, useful for typing or

PT identifying C. glutamicum or related bacteria, and as markers for

PT transformation.

XX  
XX Claim 20; Page 506-507; 1119pp; English.

XX  
XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane

CC construction and membrane transport (MCT) proteins given in AAB76510 to

CC AAB76847. The MCT nucleic acids and proteins are useful in the

CC identification of microorganisms which can be used to produce fine

SQ Sequence 177 AA;

Query Match 45.5%; Score 5; DB 16; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PKAPT 10  
 |||||  
 Db 148 pkapt 152

RESULT 37  
 AAY50043  
 ID AAY50043 standard; protein; 177 AA.  
 AC AAY50043;  
 DT 19-JAN-2000 (first entry)  
 DE Pseudomonas syringae translation initiation factor IF3.  
 KM Translation initiation factor; IF3; bacterial; screening;  
 KW antibacterial; antibiotic; drug target; initiation codon; recognition;  
 KW ubiquitous; conserved; broad spectrum; infection; contamination;  
 XX sterilisation.  
 XX Pseudomonas syringae.  
 XX WO9952357-A1.  
 PN 21-OCT-1999.  
 PD 14-APR-1999; 99WO-US08134.  
 PE 14-APR-1998; 98US-0081736.  
 PR (RIBO-) RIBOGENE INC.  
 PA Dammel GS, Watson JC, Hernandez VJ;  
 PI WPI: 1999-620242/53.  
 DR Assays for inhibitors of bacterial translation initiation factor 3 -  
 XX  
 XX  
 XX  
 PS Disclosure: Fig 1; 66pp; English.

This sequence represents Pseudomonas syringae translation initiation factor IF3. IF3 is an essential factor, acting to prevent association of ribosomal subunits and to recognise the correct initiation codon. IF3 can be used as a drug target for potential antibiotics. The invention relates to methods for screening for test compounds that inhibit the activity of IF3 which comprise using a reporter gene system in whole cells and detecting the ability of IF3 to discriminate against translation initiation at an atypical start codon of the reporter gene. IF3 is apparently ubiquitous and conserved throughout the bacterial kingdom. It is therefore likely that any compounds which are effective at inhibiting IF3 of one bacterial species will have an inhibitory effect on the IF3 of its functional equivalents in a wide range of bacteria. Importantly, IF3 has no functional homologue in mammalian cells. As a result, the potential toxicity of IF3 inhibitors is likely to be low. Compounds identified by the method that inhibit bacterial IF3 are useful for inhibiting the growth of bacteria, especially for treating an infectious disease in a human or a companion or livestock animal. The compound may be adjunctionally administered with a second antibacterial compound. IF3 inhibitors identified by the method are also useful for sterilising bacteria-contaminated objects or material. The compounds identified can penetrate and specifically kill the pathogenic bacterial cell, or arrest its growth without also adversely affecting its human, animal or plant host. This is because the target is highly selective as no homologous mammalian counterpart exists. The high-throughput primary screen allows for the easy visible identification of positive hits.

XX Sequence 177 AA;

Query Match 45.5%; Score 5; DB 20; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 Db 8 apkap 12

RESULT 38  
 AAB09006  
 ID AAB09006 standard; Protein; 177 AA.  
 AC AAB09006;  
 DT 30-AUG-2000 (first entry)  
 DE Hepatitis GB virus protein sequence SEQ ID NO:49.  
 KM Hepatitis GB virus; HGBV; diagnosis; therapeutic; immunogenic;  
 KW infection; detection; characterisation; hepatitis.  
 XX Hepatitis GB virus.  
 XX US6051374-A.  
 PN 18-APR-2000.  
 PD 07-JUN-1995; 95US-0488445.  
 PE 14-FEB-1994; 94US-0196030.  
 PR 13-MAY-1994; 94US-0242654.  
 PR 29-JUL-1994; 94US-0283314.  
 PR 23-NOV-1994; 94US-0344185.  
 PR 23-NOV-1994; 94US-0344190.  
 PR 30-JAN-1995; 95US-0377557.  
 PA (ABBO) ABBOTT LAB.  
 PI Dawson GJ, Leary TP, Muerhoff AS, Pilot-Matias TJ, Buljk SL;  
 PI Mushawar IK, Simons JN, Desai SM, Erker JC, Schlauder GG;  
 DR WPI: 2000-338307/29.  
 XX  
 XX  
 XX  
 PS Example 5; Column 171-172; 369pp; English.

The present invention describe a method for detecting target hepatitis GB virus (HGBV) nucleic acid (THN) in a test sample (T) suspected of containing HGBV. The method involves reacting (T) with a HGBV polynucleotide probe (I) containing 15 contiguous nucleotides, and which selectively hybridises to the HGBV genome or its full complement, and detecting the complex that contains THN, indicating the presence of target HGBV. The method is used for detecting target HGBV nucleic acid in the test sample suspected of containing HGBV and for characterisation of newly ascertained etiological agent of non-A, non-B, non-C, non-D and non-E hepatitis causing agents collectively termed as hepatitis GB viruses. AAY5270 to AAY5489 and AAB08985 to AAB09480 represent nucleotide and protein sequences used in the exemplification of the present invention.

SQ Sequence 177 AA;

Query Match 45.5%; Score 5; DB 21; Length 177;





CC frames in the complete genome (see AAX91990) of *Chlamydia pneumoniae*.  
 CC *C. pneumoniae* causes respiratory disease such as pneumonia and  
 CC bronchitis and is thought to be a contributing factor in heart  
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
 CC frames of the *C. pneumoniae* genome (see AAY14584-Y15879) can be used in  
 CC immunogenic compositions as vaccines. Vectors containing *C. pneumoniae*  
 CC nucleotide sequences can also be used as immunogenic compositions,  
 CC especially where the vector directs the expression of a neutralising  
 CC epitope of *C. pneumoniae*.

XX Sequence 132 AA;

Query Match 45.5%; Score 5; DB 20; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 KAPTE 11  
 |||||  
 DB 94 kapte 98

RESULT 33

AAV13180  
 ID AAV13180 standard; Protein; 135 AA.

XX AAV13180;

DE 22-JUN-1999 (first entry)

XX Human secreted protein encoded by 5' EST SEQ ID NO: 194.

XX Human; secreted protein; EST: expressed sequence tag; diagnosis:  
 KW forensic; gene therapy; chromosome mapping; signal peptide;  
 KW upstream regulatory sequence; cytokine activity; cell proliferation;  
 KW differentiation; haematopoiesis regulation; tissue growth regulation;  
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;  
 KW thrombolytic; anti-inflammatory; tumour inhibition.

OS Homo sapiens.

PN WO9906552-A2.

PD 11-FEB-1999.

PF 31-JUL-1998; 98WO-IB01236.

PR 01-AUG-1997; 97US-0905223.

PA (GEST ) GENSET.

PI Duclert A, Dumas Milne Edwards J, Lacroix B;

DR WPI; 1999-153782/13.

DR N-PSDB; AAX51980.

XX New isolated brain-derived nucleic acids - used to develop products  
 PT which may have cytokine, immune, regulatory, haematopoiesis  
 PT regulating, anti-inflammatory or tumour inhibition activity  
 XX  
 PS Claim 34; Page 556; 577pp; English.

CC AAX51787 to AAX52019 represent 5' expressed sequence tags (ESTs) for  
 CC human secreted proteins, and encode the proteins given in AAV12987 to  
 CC AAV13219, respectively. The proteins given represent the signal peptide  
 CC and an N-terminal fragment of a secreted protein. The nucleic acid  
 CC sequences can be used for producing secreted human gene products. They  
 CC can also be used to develop products for diagnosis and therapy. The  
 CC proteins obtained may have cytokine activity, cell  
 CC proliferation/differentiation activity, haematopoiesis regulating  
 CC activity, tissue growth regulating activity, reproductive hormone  
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and

CC thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, tumour inhibition activity or other activities. The products  
 CC can be used in forensic, gene therapy and chromosome mapping procedures.  
 CC The sequences can also be used for obtaining corresponding promoter  
 CC sequences. The nucleic acids encoding the signal peptide can be used for  
 CC directing extracellular secretion of a polypeptide or the insertion of a  
 CC polypeptide into a membrane, or importing a polypeptide into a cell.

SO Sequence 135 AA;

Query Match 45.5%; Score 5; DB 20; Length 135;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 QTRPK 7  
 |||||  
 DB 8 qtrpk 12

RESULT 34

AAG02253  
 ID AAG02253 standard; Protein; 157 AA.

XX AAG02253;

DE 06-OCT-2000 (first entry)

XX Human secreted protein, SEQ ID NO: 6334.

XX Human; 5' EST: expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP103401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST ) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

DR N-PSDB; AAC02259.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX  
 PS Claim 13; SEQ ID 6334; 71pp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number  
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs  
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30  
 CC different tissues. EST sequences usually correspond mainly to the 3'  
 CC untranslated region (UTR) of the mRNA because they are often obtained  
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for  
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in  
 CC those cases where longer cDNA sequences have been obtained, the full 5'  
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'  
 CC ends and can therefore be used to obtain full length cDNAs and genomic  
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and  
 CC chromosome mapping procedures. They are used to obtain upstream  
 CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 157 AA;

```

RESULT 30
AAV39038
ID AAV39038 standard; Protein: 113 AA.
XX
AC AAV39038;
XX
DT 05-NOV-1999 (first entry)
XX
DE M. tuberculosis recombinant antigen protein MO-9.
XX
KM Antigen; diagnosis; detection; infection; antibody; immunisation;
XX vaccine; immunity.
XX
OS Mycobacterium tuberculosis.
XX
PN MO9942118-A2.
XX
PD 26-AUG-1999.
XX
PF 17-FEB-1999; 99MO-US03265.
XX
PR 05-MAY-1998; 98US-0072596.
XX 18-FEB-1998; 98US-0024753.
XX
PA (CORI-) CORIXA CORP.
XX
PI Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R.
PI Lodes MJ, Reed SG, Skelky YAW, Twardzik DR, Vedvick TS;
XX
DR WPI; 1999-527416/44.
XX
PT New polypeptide comprising antigenic portions of M. tuberculosis
XX
PS Example 5; Page 260; 323pp; English.
XX
CC This invention describes novel recombinant antigens and their encoding
CC nucleic acids derived from Mycobacterium tuberculosis. The novel
CC polypeptides are useful for detecting M. tuberculosis infection in a
CC biological sample by detecting antibodies which bind with the
CC polypeptides, and are useful as vaccines for immunizing against
CC M. tuberculosis infection. The new detection methods are needed as
CC current vaccination strategies do not provide 100% immunity.
XX
SQ Sequence 113 AA;

Query Match 45.5%; Score 5; DB 20; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9
DB 103 apkap 107

RESULT 31
AAW46291
ID AAW46291 standard; protein: 119 AA.
XX
AC AAW46291;
XX
DT 29-JUL-1998 (first entry)
XX
DE Pneumococcal surface protein As (PspAs) from clade 5 strain ATCC6303.
XX
KM Streptococcus pneumoniae; vaccine; pneumococcal surface protein As;
XX infection; protection; PspAs.
XX
OS Streptococcus pneumoniae.
XX
PN MO9811915-A1.
XX

```

```

PD 26-MAR-1998.
XX
PF 22-SEP-1997; 97MO-US16761.
XX
PR 20-SEP-1996; 96US-0710749.
XX
PA (CONN-) CONNAUGHT LAB LTD.
XX
PI Becker RS, Briles DE, Hollingshead S;
XX
DR WPI; 1998-217031/19.
XX
PT New vaccines for protection against pneumococcal infection -
XX comprising at least 2 pneumococcal surface protein As, each
XX selected from a different family
XX
PS Example 3; Fig 7; 57pp; English.
XX
CC This is the sequence of a pneumococcal surface protein As (PspAs) from
XX strain ATCC6303, a representative strain of clade 5. This can be used in
XX the preparation of a vaccine composition comprising at least 2 PspAs,
XX each of which is selected from a different family. The vaccines can
XX provide for broad range protection against infection by different
XX Streptococcus pneumoniae strains.
XX
SQ Sequence 119 AA;

Query Match 45.5%; Score 5; DB 19; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8
DB 99 tapka 103

RESULT 32
AAV35711
ID AAV35711 standard; Protein: 132 AA.
XX
AC AAV35711;
XX
DT 13-SEP-1999 (first entry)
XX
DE Chlamydia pneumoniae protein not found in C. trachomatis.
XX
KM Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
XX vaccine; neutralising epitope.
XX
OS Chlamydia pneumoniae.
XX
PN WO9927105-A2.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98MO-IB01890.
XX
PR 04-NOV-1998; 98US-0107078.
XX 21-NOV-1997; 97FR-0014673.
XX
PA (GEST ) GENSET.
XX
PI Griffiths R;
XX
DR WPI; 1999-357842/30.
XX
PT Genome sequence of Chlamydia pneumoniae
XX
PS Page 1416-1417; Disclosure: 1912pp; English.
XX
AAV34584-Y35879 represent the proteins encoded by all the open reading
XX
CC

```

DR WPI: 2000-056500/05.  
 DR N-PSDB; AA240359.  
 XX  
 PT A mutant cytochrome c3 and a microbe having the mutant cytochrome c3 -  
 PT useful for improving the electron transfer reaction velocity of the heme  
 PT protein of a sulphate-reducing microbe and a photosynthetic microbe  
 XX  
 XX Example 4; Page 7; 11pp; Japanese.  
 CC This sequence is the wild type cytochrome c3 of *Desulfovibrio vulgaris*  
 CC Miyazaki F IAM2604. The invention relates to mutant cytochrome c3  
 CC proteins that are mutated near the heme 3 region, and is structurally  
 CC changed by gene manipulation so that the extent a specific amino acid  
 CC structural part controlling the electron transfer velocity of heme  
 CC protein covers heme 3 is lowered. By mutating this region of cytochrome  
 CC c3, the electron transfer reaction velocity of the heme protein of a  
 CC sulphate-reducing microbe and a photosynthetic microbe can be improved.  
 CC  
 SQ Sequence 107 AA;

Query Match 45.5%; Score 5; DB 21; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APPAP 9  
 |||||  
 DB 1 apkap 5

RESULT 28  
 AAY55100  
 ID AAY55100 standard; peptide; 107 AA.

AC AAY55100;

DT 02-MAR-2000 (first entry)

DE Mutant cytochrome c3 peptide #5.

KW Cytochrome c3; muteln; heme 3 region; electron transfer velocity;

KM heme protein; sulphate-reducing microbe; photosynthetic microbe.

OS *Desulfovibrio vulgaris*.

OS Synthetic.

PM JP11290081-A.

PD 26-OCT-1999.

PF 15-APR-1998; 96JP-0104667.

PR 15-APR-1998; 96JP-0104667.

PA (MITO) MITSUBISHI JUKOGYO KK.

DR WPI: 2000-056500/05.

XX A mutant cytochrome c3 and a microbe having the mutant cytochrome c3 -  
 PT useful for improving the electron transfer reaction velocity of the heme  
 PT protein of a sulphate-reducing microbe and a photosynthetic microbe  
 XX  
 XX Example 5; Page 9; 11pp; Japanese.

CC This sequence represents a mutant cytochrome c3 of the invention. It is  
 CC a mutant of the *Desulfovibrio vulgaris* Hildenborough ATCC29579 sequence.  
 CC The protein is mutated near the heme 3 region, and is structurally  
 CC changed by gene manipulation so that the extent a specific amino acid  
 CC structural part controlling the electron transfer velocity of heme  
 CC protein covers heme 3 is lowered. By mutating this region of cytochrome  
 CC c3, the electron transfer reaction velocity of the heme protein of a  
 CC sulphate-reducing microbe and a photosynthetic microbe can be improved.  
 CC

SQ Sequence 107 AA;

Query Match 45.5%; Score 5; DB 21; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APPAP 9  
 |||||  
 DB 1 apkap 5

RESULT 29  
 AAY39181  
 ID AAY39181 standard; Protein; 113 AA.

AC AAY39181;

DT 05-NOV-1999 (first entry)

DE M. tuberculosis antigen 5' MO-9 amino acid sequence.

KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;

KM immunotherapy; diagnosis; immunisation; vaccine; infection;

KW immune response; skin test.

OS Mycobacterium tuberculosis.

PN WO9942076-A2.

PD 26-AUG-1999.

PF 17-FEB-1999; 99WO-US03268.

PR 05-MAY-1998; 98US-0072967.

PR 18-FEB-1998; 98US-0025197.

PA (CORI) CORIXA CORP.

PI Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;

PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;

DR WPI: 1999-527409/44.

DR N-PSDB; AA219373.

PT New antigens from Mycobacterium tuberculosis useful in diagnostic  
 PT skin tests and protective or therapeutic vaccines or compositions

PS Example 5; Page 215; 299pp; English.

CC The present invention describes polypeptides comprising an immunogenic  
 CC part of a Mycobacterium tuberculosis antigen (Ag). Also described  
 CC are vaccines and fusion protein containing M. tuberculosis Ag's.  
 CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and  
 CC other polypeptides fragments, can be used in pharmaceutical compositions  
 CC or vaccines to generate a protective or therapeutic immune response to  
 CC M. tuberculosis and as reagents in skin tests for diagnosis of  
 CC tuberculosis. Ag can induce proliferation of, or cytokine secretion  
 CC by, T, B or natural killer cells and/or macrophages in  
 CC tuberculosis-immune subjects. AA219249 to AA219460 and AAY39083 to  
 CC AAY39225 are used in the exemplification of the present invention.  
 CC

SQ Sequence 113 AA;

Query Match 45.5%; Score 5; DB 20; Length 113;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APPAP 9  
 |||||  
 DB 103 apkap 107

Query Match 45.5%: Score 5; DB 21; Length 107;

Best Local Similarity 100.0%; Pred. No. 1.4e+02; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
| | | | |  
Db 1 apkap 5

RESULT 25

AAV55097 ID AAV55097 standard; peptide; 107 AA.

XX AAV55097;

DT 02-MAR-2000 (first entry)

DE Mutant cytochrome c3 peptide #3.

KM Cytochrome c3; mutcin; heme 3 region; electron transfer velocity;

KW heme protein; sulphate-reducing microbe; photosynthetic microbe.

OS Desulfovibrio vulgaris.

XX Synthetic.

PN JP11290081-A.

PD 26-OCT-1999.

PF 15-APR-1998; 98JP-0104667.

PR 15-APR-1998; 98JP-0104667.

PA (MITO ) MITSUBISHI JUKOGYO KK.

DR WPI; 2000-056500/05.

PT A mutant cytochrome c3 and a microbe having the mutant cytochrome c3 -

PT useful for improving the electron transfer reaction velocity of the heme

PS protein of a sulphate-reducing microbe and a photosynthetic microbe

XX Claim 5; Page 9; 11pp; Japanese.

CC This sequence represents a mutant cytochrome c3 of the invention. It is

CC a mutant of the Desulfovibrio vulgaris Hildenborough ATCC29579 sequence.

CC The protein is mutated near the heme 3 region, and is structurally

CC changed by gene manipulation so that the extent a specific amino acid

CC structural part controlling the electron transfer velocity of heme

CC protein covers heme 3 is lowered. By mutating this region of cytochrome

CC c3, the electron transfer reaction velocity of the heme protein of a

CC sulphate-reducing microbe and a photosynthetic microbe can be improved.

XX Sequence 107 AA;

OY 5 APKAP 9  
| | | | |  
Db 1 apkap 5

RESULT 26

AAV55098 ID AAV55098 standard; peptide; 107 AA.

XX AAV55098;

DT 02-MAR-2000 (first entry)

XX Mutant cytochrome c3 peptide #4.

KW Cytochrome c3; mutcin; heme 3 region; electron transfer velocity;

KW heme protein; sulphate-reducing microbe; photosynthetic microbe.

OS Desulfovibrio vulgaris.

XX Synthetic.

PN JP11290081-A.

PD 26-OCT-1999.

PF 15-APR-1998; 98JP-0104667.

PR 15-APR-1998; 98JP-0104667.

PA (MITO ) MITSUBISHI JUKOGYO KK.

DR WPI; 2000-056500/05.

PT A mutant cytochrome c3 and a microbe having the mutant cytochrome c3 -

PT useful for improving the electron transfer reaction velocity of the heme

PS protein of a sulphate-reducing microbe and a photosynthetic microbe

XX Claim 6; Page 9-10; 11pp; Japanese.

CC This sequence represents a mutant cytochrome c3 of the invention. It is

CC a mutant of the Desulfovibrio vulgaris Hildenborough ATCC29579 sequence.

CC The protein is mutated near the heme 3 region, and is structurally

CC changed by gene manipulation so that the extent a specific amino acid

CC structural part controlling the electron transfer velocity of heme

CC protein covers heme 3 is lowered. By mutating this region of cytochrome

CC c3, the electron transfer reaction velocity of the heme protein of a

CC sulphate-reducing microbe and a photosynthetic microbe can be improved.

XX Sequence 107 AA;

OY 5 APKAP 9  
| | | | |  
Db 1 apkap 5

RESULT 27

AAV55099 ID AAV55099 standard; Protein; 107 AA.

XX AAV55099;

DT 02-MAR-2000 (first entry)

DE Wild type cytochrome c3 protein sequence.

KW Cytochrome c3; mutcin; heme 3 region; electron transfer velocity;

KW heme protein; sulphate-reducing microbe; photosynthetic microbe.

OS Desulfovibrio vulgaris.

PN JP11290081-A.

PD 26-OCT-1999.

PF 15-APR-1998; 98JP-0104667.

PR 15-APR-1998; 98JP-0104667.

PA (MITO ) MITSUBISHI JUKOGYO KK.

KW Cytochrome c3; sulfate-reducing microbe; mutant; bio-element;  
 KW molecular wire; organic semiconductor; commutator element.  
 OS Desulfovibrio vulgaris.  
 XX JP2000217577-A.  
 PN 08-AUG-2000.  
 XX 28-JAN-1999; 99JP-0020018.  
 XX 28-JAN-1999; 99JP-0020018.  
 XX (MITO ) MITSUBISHI JUKOGYO KK.  
 PA WPI: 2000-605045/58.  
 DR N-PSDB: AAA98362.  
 XX  
 XX Production of mutant cytochrome C3, useful as a material for a bio  
 PT element such as an organic semiconductor, by point-mutating wild type  
 PT cytochrome C3 from a sulfate reducing microbe -  
 PS Claim 1; Page 8; 9pp; Japanese.  
 XX  
 XX This invention describes the novel production of a mutant cytochrome C3  
 CC comprising a sequence (A) of 107 amino acids (fully defined in the  
 CC specification), by point-mutating a wild type cytochrome C3 derived from  
 CC a sulfate reducing microbe. The new method includes the following steps:  
 CC (1) a plasmid carrying a gene of the mutant cytochrome C3 is introduced  
 CC to an aerobic microbe to transform the aerobic microbe; (2) culturing  
 CC aerobically the transformed aerobic microbe to grow it; and (3)  
 CC collecting the mutant cytochrome C3 produced by the transformed aerobic  
 CC microbe. Cytochrome C3 can be used as a material for a bio element such  
 CC as a molecular wire, organic semiconductor and a commutator element.  
 CC  
 SQ Sequence 107 AA:

Query Match 45.5%; Score 5; DB 21; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APPAP 9  
 I I I I I  
 DB 1 appap 5

RESULT 23  
 AAY55095  
 ID AAY55095 standard; peptide: 107 AA.  
 XX  
 AC AAY55095;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE Mutant cytochrome c3 peptide #1.  
 XX  
 KW Cytochrome c3; muteln; heme 3 region; electron transfer velocity;  
 KW heme protein; sulphate-reducing microbe; photosynthetic microbe.  
 XX  
 OS Desulfovibrio vulgaris.  
 OS Synthetic.  
 XX  
 PN JP11290081-A.  
 XX  
 PD 26-OCT-1999.  
 XX  
 PF 15-APR-1998; 98JP-0104667.  
 XX  
 PR 15-APR-1998; 98JP-0104667.  
 XX  
 PA (MITO ) MITSUBISHI JUKOGYO KK.  
 XX

DR WPI: 2000-056500/05.  
 XX  
 XX A mutant cytochrome c3 and a microbe having the mutant cytochrome c3 -  
 PT useful for improving the electron transfer reaction velocity of the heme  
 PT protein of a sulphate-reducing microbe and a photosynthetic microbe  
 XX  
 XX Claim 2; Page 8; 11pp; Japanese.  
 XX  
 XX This sequence represents a mutant cytochrome c3 of the invention. It is  
 CC a mutant of the Desulfovibrio vulgaris Miyazaki F IAM2604 sequence.  
 CC The protein is mutated near the heme 3 region, and is structurally  
 CC changed by gene manipulation so that the extent a specific amino acid  
 CC structural part controlling the electron transfer velocity of heme  
 CC protein covers heme 3 is lowered. By mutating this region of cytochrome  
 CC c3, the electron transfer reaction velocity of the heme protein of a  
 CC sulphate-reducing microbe and a photosynthetic microbe can be improved.  
 CC  
 SQ Sequence 107 AA:

Query Match 45.5%; Score 5; DB 21; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APPAP 9  
 I I I I I  
 DB 1 appap 5

RESULT 24  
 AAY55096  
 ID AAY55096 standard; peptide: 107 AA.  
 XX  
 AC AAY55096;  
 XX  
 DT 02-MAR-2000 (first entry)  
 XX  
 DE Mutant cytochrome c3 peptide #2.  
 XX  
 KW Cytochrome c3; muteln; heme 3 region; electron transfer velocity;  
 KW heme protein; sulphate-reducing microbe; photosynthetic microbe.  
 XX  
 OS Desulfovibrio vulgaris.  
 OS Synthetic.  
 XX  
 PN JP11290081-A.  
 XX  
 PD 26-OCT-1999.  
 XX  
 PF 15-APR-1998; 98JP-0104667.  
 XX  
 PR 15-APR-1998; 98JP-0104667.  
 XX  
 PA (MITO ) MITSUBISHI JUKOGYO KK.  
 XX  
 DR WPI: 2000-056500/05.  
 XX  
 PT A mutant cytochrome c3 and a microbe having the mutant cytochrome c3 -  
 PT useful for improving the electron transfer reaction velocity of the heme  
 PT protein of a sulphate-reducing microbe and a photosynthetic microbe  
 XX  
 XX Claim 3; Page 8; 11pp; Japanese.  
 XX  
 XX This sequence represents a mutant cytochrome c3 of the invention. It is  
 CC a mutant of the Desulfovibrio vulgaris Miyazaki F IAM2604 sequence.  
 CC The protein is mutated near the heme 3 region, and is structurally  
 CC changed by gene manipulation so that the extent a specific amino acid  
 CC structural part controlling the electron transfer velocity of heme  
 CC protein covers heme 3 is lowered. By mutating this region of cytochrome  
 CC c3, the electron transfer reaction velocity of the heme protein of a  
 CC sulphate-reducing microbe and a photosynthetic microbe can be improved.  
 CC  
 SQ Sequence 107 AA:

XX The present sequence represents a cytochrome c3 protein used in an  
 CC example of the present invention. A new method has been developed  
 CC for improving electron moving velocity (EMV) of haem protein. The  
 CC method comprises controlling in the specific aa structure EMV of  
 CC haem protein around the haem 3 of cytochrome c3, where the degree of  
 CC the aa structure covering haem 3 is lowered. Also described in the  
 CC present invention is a method for improving the reaction velocity of  
 CC redox reaction of a sulphate reducing microorganism having cytochrome c3  
 CC performed analogically to the method above. The velocity and the yield  
 CC of bioreaction can efficiently be improved.  
 XX  
 SQ Sequence 107 AA;

Query Match 45.5%; Score 5; DB 19; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 Db 1 apkap 5

RESULT 20  
 AAM57221  
 ID AAM57221 standard; Protein; 107 AA.

AC AAM57221;  
 XX  
 DT 04-AUG-1998 (first entry)  
 XX

DE Cytochrome c3 protein SEQ ID NO:5 from Desulfovibrio vulgaris.

KM Desulfovibrio vulgaris Miyazaki F IAM2604; haem 3; cytochrome c3;  
 KM electron moving velocity; haem protein; heme.

OS Desulfovibrio vulgaris.

XX JP10113182-A.

PD 06-MAY-1998.

PF 11-OCT-1996; 96JP-0269698.

PR 11-OCT-1996; 96JP-0269698.

XX (MITO ) MITSUBISHI JUKOGYO KK.

XX WPI; 1998-315472/28.

PT Improvement of electron moving velocity of haem protein - by  
 PT controlling degree of amino acid structure covering haem protein

PS Example 5; Page 8; 11pp; Japanese.

XX The present sequence represents a cytochrome c3 protein used in an  
 CC example of the present invention. A new method has been developed  
 CC for improving electron moving velocity (EMV) of haem protein. The  
 CC method comprises controlling in the specific aa structure EMV of  
 CC haem protein around the haem 3 of cytochrome c3, where the degree of  
 CC the aa structure covering haem 3 is lowered. Also described in the  
 CC present invention is a method for improving the reaction velocity of  
 CC redox reaction of a sulphate reducing microorganism having cytochrome c3  
 CC performed analogically to the method above. The velocity and the yield  
 CC of bioreaction can efficiently be improved.  
 XX  
 SQ Sequence 107 AA;

Query Match 45.5%; Score 5; DB 19; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 Db 1 apkap 5

RESULT 21  
 AAB10915  
 ID AAB10915 standard; Protein; 107 AA.

XX AAB10915;

DT 01-FEB-2001 (first entry)

DE D. vulgaris cytochrome C3 protein.

KM Cytochrome C3; sulfate-reducing microbe; mutant; bio-element;  
 KM molecular wire; organic semiconductor; commutator element.

OS Desulfovibrio vulgaris.

XX JP2000217577-A.

PD 08-AUG-2000.

PF 28-JAN-1999; 99JP-0020018.

PR 28-JAN-1999; 99JP-0020018.

XX (MITO ) MITSUBISHI JUKOGYO KK.

XX WPI; 2000-605045/58.

DR N-PSDB; AAA98361.

XX Production of mutant cytochrome C3, useful as a material for a bio  
 PT element such as an organic semiconductor, by point-mutating wild type  
 PT cytochrome C3 from a sulfate reducing microbe -

XX Disclosure; Page 6; 9pp; Japanese.

PS This invention describes the novel production of a mutant cytochrome C3  
 CC comprising a sequence (A) of 107 amino acids (fully defined in the  
 CC specification), by point-mutating a wild type cytochrome C3 derived from  
 CC a sulfate reducing microbe. The new method includes the following steps:  
 CC (1) a plasmid carrying a gene of the mutant cytochrome C3 is introduced  
 CC to an aerobic microbe to transform the aerobic microbe; (2) culturing  
 CC aerobically the transformed aerobic microbe to grow it; and (3)  
 CC collecting the mutant cytochrome C3 produced by the transformed aerobic  
 CC microbe. Cytochrome C3 can be used as a material for a bio element such  
 CC as a molecular wire, organic semiconductor and a commutator element.  
 XX

SQ Sequence 107 AA;

Query Match 45.5%; Score 5; DB 21; Length 107;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9  
 |||||  
 Db 1 apkap 5

RESULT 22

AAB10916  
 ID AAB10916 standard; Protein; 107 AA.

XX AAB10916;

DT 01-FEB-2001 (first entry)

DE D. vulgaris cytochrome C3 mutant protein.

XX

```

RESULT 17
AAW57218
ID AAW57218 standard; Protein; 107 AA.
XX
AC AAW57218;
XX
DT 04-AUG-1998 (first entry)
XX
DE Desulfovibrio vulgaris Miyazaki F IAM2604 haem 3 of cytochrome c3.
XX
KM Desulfovibrio vulgaris Miyazaki F IAM2604; haem 3; cytochrome c3;
XX electron moving velocity; haem protein; heme.
XX
OS Desulfovibrio vulgaris.
XX
PN JP10113182-A.
XX
PD 06-MAY-1998.
XX
PF 11-OCT-1996; 96JP-0269698.
XX
PR 11-OCT-1996; 96JP-0269698.
XX
PA (MITO ) MITSUBISHI JUKOGYO KK.
XX
DR WPI; 1998-315472/28.
XX N-PSDB; AAW28847.
XX
PT Improvement of electron moving velocity of haem protein - by
XX controlling degree of amino acid structure covering haem protein
XX
PS Claim 3; Page 6; 11pp; Japanese.
XX
CC The present sequence represents haem 3 from cytochrome c3 of
CC Desulfovibrio vulgaris Miyazaki F IAM2604. A new method has been
CC developed for improving electron moving velocity (EMV) of haem protein.
CC The method comprises controlling in the specific aa structure EMV of
CC haem protein around the haem 3 of cytochrome c3, where the degree of
CC the aa structure covering haem 3 is lowered. Also described in the
CC present invention is a method for improving the reaction velocity of
CC redox reaction of a sulphate reducing microorganism having cytochrome c3
CC performed analogically to the method above. The velocity and the yield
CC of bioreaction can efficiently be improved.
XX
SQ Sequence 107 AA;

Query Match 45.5%; Score 5; DB 19; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9
DB 1 apkap 5

RESULT 18
AAW57219
ID AAW57219 standard; Protein; 107 AA.
XX
AC AAW57219;
XX
DT 04-AUG-1998 (first entry)
XX
DE Cytochrome c3 protein SEQ ID NO:3 from J10113182 Example 1.
XX
KM Desulfovibrio vulgaris Miyazaki F IAM2604; haem 3; cytochrome c3;
XX electron moving velocity; haem protein; heme.
XX
OS Synthetic.
XX Desulfovibrio vulgaris.

```

```

XX
PN JP10113182-A.
XX
PD 06-MAY-1998.
XX
PF 11-OCT-1996; 96JP-0269698.
XX
PR 11-OCT-1996; 96JP-0269698.
XX
PA (MITO ) MITSUBISHI JUKOGYO KK.
XX
DR WPI; 1998-315472/28.
XX
PT Improvement of electron moving velocity of haem protein - by
XX controlling degree of amino acid structure covering haem protein
XX
PS Example 1; Page 7; 11pp; Japanese.
XX
CC The present sequence represents a cytochrome c3 protein used in an
CC example of the present invention. A new method has been developed
CC for improving electron moving velocity (EMV) of haem protein. The
CC method comprises controlling in the specific aa structure EMV of
CC haem protein around the haem 3 of cytochrome c3, where the degree of
CC the aa structure covering haem 3 is lowered. Also described in the
CC present invention is a method for improving the reaction velocity of
CC redox reaction of a sulphate reducing microorganism having cytochrome c3
CC performed analogically to the method above. The velocity and the yield
CC of bioreaction can efficiently be improved.
XX
SQ Sequence 107 AA;

Query Match 45.5%; Score 5; DB 19; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APKAP 9
DB 1 apkap 5

RESULT 19
AAW57220
ID AAW57220 standard; Protein; 107 AA.
XX
AC AAW57220;
XX
DT 04-AUG-1998 (first entry)
XX
DE Cytochrome c3 protein SEQ ID NO:4 from J10113182 Example 1.
XX
KM Desulfovibrio vulgaris Miyazaki F IAM2604; haem 3; cytochrome c3;
XX electron moving velocity; haem protein; heme.
XX
OS Synthetic.
XX Desulfovibrio vulgaris.
XX
PN JP10113182-A.
XX
PD 06-MAY-1998.
XX
PF 11-OCT-1996; 96JP-0269698.
XX
PR 11-OCT-1996; 96JP-0269698.
XX
PA (MITO ) MITSUBISHI JUKOGYO KK.
XX
DR WPI; 1998-315472/28.
XX
PT Improvement of electron moving velocity of haem protein - by
XX controlling degree of amino acid structure covering haem protein
XX
PS Example 1; Page 7-8; 11pp; Japanese.

```



CC anticancer and cytotoxic activity. (A) are used (1) for recombinant  
CC expression of polypeptides (B) and (11) to isolate complete genes. (B)  
CC are used (1) to identify agents suitable for treatment of uterine or  
CC endometrial cancer; (11) directly for treating these forms of cancer  
CC (including expression from gene therapy vectors) and (11) for  
CC generation of specific antibodies. (A) are identified by assembling ESTs  
CC (expressed sequence tags) from a particular tissue type before comparison  
CC of expression patterns. This allows a significantly longer fragment of  
CC the gene to be revealed, so should reduce the number of failures  
CC associated with the fact that ESTs from different libraries may represent  
CC different parts of the same unknown gene, distorting the estimated  
CC frequency of occurrence in a particular tissue. AY559941-Y60328 represent  
CC protein fragments encoded by the human endometrium tumour cDNA library  
CC derived EST fragments represented in AA415981-742121.

Query Match	45.5%	Score 5	DB 20	Length 95
Best local similarity	100.0%	Pred. No.	1.2e+02	
Matches	5	Conservative	0	Mismatches 0
				Indels 0
				Gaps 0

QY	7	KAPTE	11
Db	21	kapte	25

**RESULT 15**

AAY29500  
ID AAY29500 standard; Protein; 96 AA.

DT 13-OCT-1999 (first entry)

DE Human lung tumour protein LT86-27 predicted amino acid sequence.

KW Human; lung tumour protein; diagnosis; lung cancer; vaccine;

OS Homo sapiens.

PD 05-AUG-1999.

PF 26-JAN-1999; 99WO-US01642.

PR 22-DEC-1998; 98US-0219245.

PR . 28-JAN-1998; 98US-0015029.  
PR 18-MAR-1998; 98US-0040838

PR	18-MAR-1998;	98US-0040831.
PR	23-JUL-1998;	98US-0122191

PR 43-JUL-1998; 9805-0122192-  
XX

PA (CORP-) CORIAX CORP.  
XX

FI  
FI  
XX

DR N-PSDB; AA207170.

PT Lung tumour specific polynucleotides for inhibiting the development of lung cancer

PS Example 2; Page 99; 171pp; English.

The present invention describes a method for

CC represent specifically claimed polynucleotides, and AAY29486 to AAY29571

CC pharmaceutical compositions and vaccines to inhibit the development of  
CC lung cancer. They can also be used to detect lung cancer in a patient.  
CC Probes and antibodies derived from the lung tumour sequences are useful  
CC in detection of lung cancer.  
CC

Query Match	45.5%	Score 5:	DB 20:	Length 96:
Best Local Similarity	100.0%	Pred. No.	1.2e+02:	
Matches	5:	Conservative	0:	Indels 0:
		Mismatches	0:	Gaps 0

```
QY      6 PKAPT 10
        |||||
Db      .36 pkapt 40
```

## RESULT 16

AAB44435  
ID AAB44435 standard; Protein; 96 AA.

DT 05-FEB-2001 (first entry)

Human lung tumour-specific antigen encoded by cDNA #46.

**KW** Lung tumour protein; lung cancer; cytostatic; vaccine.

OS Homo sapiens.

PD 12-OCT-2000

PF 30-MAR-2000; 2000WO-US08560;  
XX

PR	02-APR-1999;	99US-0285323.
PR	09-AUG-1999:	99US-0370838.

PR 30-DEC-1999; 9905-04/6233  
PR 03-MAR-2000: 2000US-0518809

AA  
PA (CORI-) CORIXA CORP.

Reed SG, Lodes MJ, Mohamat

DR WPI; 2000-638466/61.

XX

PT detecting, monitoring

PS Example 2; Page 120; 243pp; English.

CC The present sequence is given in a

The present sequence is given in a specification relating to compounds for therapy and diagnosis of lung cancer. Polypeptides comprising at least an immunogenic part of a lung tumour protein are disclosed. The polypeptides are useful for inhibiting the development of cancer, especially lung cancer. Samples of T cells expressing the polypeptides may be used to inhibit the development of cancer. The polypeptides are also useful for detecting and monitoring the progression of cancer, especially lung cancer.

50 Sequence 96 AA;

Query Match	45.5%	Score 5:	DB 21:	length 96:
Best Local Similarity	100.0%	Pred. No.	1.2e+02:	
Matches	5:	Conservative	0:	Indels 0:
		Mismatches	0:	Gaps 0

QY	6	PKAPT	10
Db	36	pkapt	40

CC which may be used in the prodn. of antibodies for the detection  
 CC of *Listeria* by immunoassay (partic. ELISA). It may be used as  
 CC part of a method that allows determination of individual *Listeria*  
 CC species, esp. *L. monocytogenes*.  
 CC  
 XX

SQ Sequence 25 AA:

Query Match 45.5%; Score 5; DB 14; Length 25;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QOTAP 6  
 |||||  
 Db 14 qctap 18

RESULT 13

AA94958  
 ID AAY94958 standard; Protein; 54 AA.

AC AAY94958;

DT 16-JUN-2000 (first entry)

DE Human secreted protein clone LR9\_3 protein sequence SEQ ID NO:122.

XX Human; secreted protein; immunostimulant; immunosuppressant; virucide;  
 XX antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;  
 XX antidiabetic; antiaesthetic; antirheumatic; antiprotozoal;  
 XX antihypertensive; immune deficiency; severe combined immunodeficiency; SCID;  
 XX infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;  
 XX connective tissue disease; multiple sclerosis; erythematosis;  
 XX rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;  
 XX Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;  
 XX insulin dependent diabetes mellitus; graft-versus-host-disease;  
 XX autoimmune inflammatory eye disease; allergy.

OS Homo sapiens.

PN WO200009552-A1.

PD 24-FEB-2000.

PF 13-AUG-1999; 99WO-US18298.

PR 14-AUG-1998; 98US-0096622.

PR 17-AUG-1998; 98US-0096815.

PR 04-SEP-1998; 98US-0099229.

PR 23-OCT-1998; 98US-0105368.

PR 08-JAN-1999; 99US-0115234.

PR 12-FEB-1999; 99US-0119931.

PR 18-FEB-1999; 99US-0120575.

PR 30-APR-1999; 99US-0132020.

PR 11-AUG-1999; 99US-0096622.

PA (GENY ) GENETICS INST INC.

PI Jacobs K, McCoy JM, Lavallee ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M, Agostino M, Steinfinger RJ, Spaulding V;

PI Wong GG, Clark HF, Fechtel K;

DR WPI: 2000-205979/18.

XX New polynucleotides encoding secreted proteins, which may have e.g.

PT nutritional, chemokine, immune stimulating or suppressing,

PT hematopoiesis regulating, tissue growth, activin/inhibin

PT antiinflammatory or tumor inhibition activity  
 XX Claim 131; Page 581; 641pp; English.  
 PS  
 XX AAA1618 to AAA1697 encoded the human secreted proteins given in  
 CC AAY94898 to AAY94980, isolated from human adult brain, adult thyroid,

CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,  
 CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,  
 CC foetal brain, adult thymus, foetal placenta, adult uterus, adult tumour,  
 CC and adult bladder. cDNA libraries. The polynucleotides and proteins are  
 CC predicted to have biological activities which would make them suitable  
 CC for treating, preventing or ameliorating medical conditions in humans  
 CC and animals. The polynucleotides can be used as markers for tissues in  
 CC which the protein is preferentially expressed, as molecular weight  
 CC markers on Southern gels, and as chromosome markers or tags to identify  
 CC chromosomes or to map gene positions. The proteins can be used in the  
 CC treatment of immune deficiencies and disorders, such as severe combined  
 CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other  
 CC infections. These infections include human immunodeficiency virus (HIV),  
 CC hepatitis, herpesviruses, mycobacteria, *Leishmania* spp., malaria and  
 CC candidiasis. The proteins can be used to treat autoimmune disorders such  
 CC as connective tissue disease, multiple sclerosis, systemic lupus  
 CC erythematosis, rheumatoid arthritis, autoimmune pulmonary inflammation,  
 CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent  
 CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and  
 CC autoimmune inflammatory eye disease. The proteins can also be used to  
 CC treat allergic conditions, such as asthma. AAA16698 to AAA16774 represent  
 CC probes for the human secreted proteins from the present invention.  
 CC  
 XX

SQ Sequence 54 AA:

Query Match 45.5%; Score 5; DB 21; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 TAPKA 8  
 |||||  
 Db 45 tapka 49

RESULT 14

AA60235  
 ID AAY60235 standard; Protein; 95 AA.

AC AAY60235;

DT 31-JAN-2000 (first entry)

DE Human endometrium tumour EST encoded protein 295.

XX Endometrium; human; tumour; cancer; antineoplastic; cytostatic; EST;  
 XX treatment; uterine; gene therapy; expressed sequence tag.

OS Homo sapiens.

PN DE19817948-A1.

PD 21-OCT-1999.

PF 17-APR-1998; 98DE-1017948.

PR 17-APR-1998; 98DE-1017948.

PR (META-) METAGEN GES GENOMFORSCHUNG MBH.

PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;

PI WPI: 1999-591957/51.

DR N-PSDB; AA242076.

XX New nucleic acid sequences expressed in uterine cancer tissues, and  
 XX derived polypeptides, for treatment of uterine and endometrial cancer  
 XX and identification of therapeutic agents -  
 XX Claim 23; Page 392; 444pp; German.

PS  
 XX This invention describes novel human nucleic acid (cDNA) sequences (A),  
 CC that are highly expressed in uterine tumour tissue and which have

```

XX 16-JUN-1994 (first entry)
XX Listeria p60 peptide epitope.
DE Listeria monocytogenes; antibodies; immunoassay; conjugate.
XX Synthetic.
XX DE4318450-A.
XX 16-DEC-1993.
XX 03-JUN-1993; 93DE-4318450.
XX 11-JUN-1992; 92DE-4219111.
XX 25-NOV-1992; 92DE-4239567.
XX (MERE ) MERCK PATENT GMBH.
XX Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;
XX Hofmann G, Bubert A, Goebel W, Koehler S;
XX WPI; 1993-406956/51.
XX New primers for PCR detection of Listeria - including individual
XX species, also new peptide(s) for raising antibodies for
XX immunochemical detection
XX Disclosure; Page 3; 19pp; German.
XX The sequence is that of a Listeria p60 peptide epitope which
XX which may be used in the prodn. of antibodies for the detection
XX of Listeria by immunoassay (partic. ELISA). It may be used as
XX part of a method that allows determination of individual Listeria
XX species, esp. L. monocytogenes.
XX Sequence 11 AA;

Query Match 45.58; Score 5; DB 14; Length 11;
Best Local Similarity 100.08; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QOTAP 6
DB 7 qctap 11

RESULT 11
AAR54625
ID AAR54625 standard; Protein; 13 AA.
XX AAR54625;
AC 16-JUN-1994 (first entry)
XX 16-JUN-1994 (first entry)
XX Listeria innocua p60 peptide epitope.
XX Antibodies; immunoassay; conjugate.
XX Synthetic.
XX DE4318450-A.
XX 16-DEC-1993.
XX 03-JUN-1993; 93DE-4318450.
XX 11-JUN-1992; 92DE-4219111.
XX 25-NOV-1992; 92DE-4239567.
XX (MERE ) MERCK PATENT GMBH.

```

```

XX Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;
XX Hofmann G, Bubert A, Goebel W, Koehler S;
XX WPI; 1993-406956/51.
XX New primers for PCR detection of Listeria - including individual
XX species, also new peptide(s) for raising antibodies for
XX immunochemical detection
XX Disclosure; Fig 5; 19pp; German.
XX The sequence is that of a Listeria innocua p60 peptide epitope which
XX which may be used in the prodn. of antibodies for the detection
XX of Listeria by immunoassay (partic. ELISA). It may be used as
XX part of a method that allows determination of individual Listeria
XX species.
XX Sequence 13 AA;

Query Match 45.58; Score 5; DB 14; Length 13;
Best Local Similarity 100.08; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QOTAP 6
DB 8 qctap 12

RESULT 12
AAR45165
ID AAR45165 standard; Protein; 25 AA.
XX AAR45165;
AC 16-JUN-1994 (first entry)
XX 16-JUN-1994 (first entry)
XX Listeria p60 peptide epitope.
XX Listeria monocytogenes; antibodies; immunoassay; conjugate.
XX Synthetic.
XX Key Location/Qualifiers
XX Region /note="opt. 0-7 amino acid residues"
XX FT 19..25
XX FT /note="opt. 0-7 amino acid residues"
XX PN DE4318450-A.
XX 16-DEC-1993.
XX 03-JUN-1993; 93DE-4318450.
XX 11-JUN-1992; 92DE-4219111.
XX 25-NOV-1992; 92DE-4239567.
XX (MERE ) MERCK PATENT GMBH.
XX Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;
XX Hofmann G, Bubert A, Goebel W, Koehler S;
XX WPI; 1993-406956/51.
XX New primers for PCR detection of Listeria - including individual
XX species, also new peptide(s) for raising antibodies for
XX immunochemical detection
XX Claim 3; Page 4; 19pp; German.
XX The sequence is that of a Listeria p60 peptide epitope which

```

Query Match 63.6%; Score 7; DB 16; Length 7;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 OOTAPKA 8  
 1111111  
 Db 1 qqlapka 7

## RESULT 8

AAV26181  
 ID AAV26181 standard; Protein; 150 AA.

AC AAV26181;

DT 22-OCT-1999 (first entry)

XX Rice acetohydroxyacid reductoisomerase.

DE Acetohydroxyacid reductoisomerase; rlr24.pk0087.e7 clone; rice; BCABE;  
 KW branched chain biosynthetic enzyme; herbicide; immunological screening;  
 KM probe; inhibitor; plant breeding; gene mapping.

XX Oryza sativa.

OS Oryza sativa.

XX Key Location/Qualifiers

FT Misc-difference 14 /note= "encoded by AAC"

FT Misc-difference 16 /note= "encoded by NCC"

FT Misc-difference 86 /note= "encoded by GAN"

FT Misc-difference 108 /note= "encoded by AAN"

FT Misc-difference 118 /note= "encoded by CAN"

PN WO928446-A2.

XX 10-JUN-1999.

XX 18-NOV-1998; 98WO-US24615.

XX 02-DEC-1997; 97US-0067388.

XX (DUPO ) DU PONT DE NEMOURS & CO E. I.

XX Allen SM, Falco SC, Hiltz WD;

XX WPI: 1999-385378/32.

XX N-PSDB: AAX80619.

XX New isolated branched chain amino acid biosynthetic genes

PS Claim 6; Page 48-49; 69pp; English.

CC The present sequence is a portion of a rice acetohydroxyacid  
 CC reductoisomerase, which is a branched chain amino acid biosynthetic  
 CC enzyme (BCABE). The enzyme is encoded by a partial cDNA insert in clone  
 CC rlr24.pk0087.e7. The clone is derived from rice leaf 15 days after  
 CC germination, 24 hours after infection of Magaporthe grisea 4360-R-62  
 CC strain. Host cells, transformed by a chimeric gene comprising the  
 CC nucleic acid fragment, are used for the production of BCABE. The enzyme  
 CC is involved in biosynthesis and utilization of branched chain amino  
 CC acids, such as valine, isoleucine, and leucine. The products can be  
 CC used for altering the levels of branched chain aminoacids in plants.  
 CC They can also be used as targets to facilitate identification and/or  
 CC design of inhibitors of those enzymes that may be useful as herbicides.  
 CC The nucleic acid fragments can also be used as probes for mapping the  
 CC genes that they are a part of, and as markers for traits linked to those  
 CC genes. Such information may be useful in plant breeding in order to  
 CC develop lines with desired phenotypes.

XX SQ Sequence 150 AA;

Query Match 54.5%; Score 6; DB 20; Length 150;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 APRAPT 10  
 1111111  
 Db 22 aprapt 27

## RESULT 9

AAV26181  
 ID AAV26181 standard; Protein; 890 AA.

AC AAV26181;

DT 15-FEB-1995 (first entry)

XX NF-ATP.

DE NF-ATP; NF-ATP; preexisting nuclear factor of activated T-cell protein;

KW T-lymphocyte; immunosuppressive.

XX Mus sp.

PN WO9415964-A.

XX 21-JUL-1994.

XX 18-JAN-1994; 94WO-US00545.

XX 15-JAN-1993; 93US-0006067.

XX 11-FEB-1993; 93US-0017052.

XX 29-OCT-1993; 93US-0145006.

XX (DAND ) DANA FARBER CANCER INST INC.

XX (HARD ) HARVARD COLLEGE.

XX Hogan PG, Jain J, Mc Caffrey P, Rao A;

XX WPI: 1994-249140/30.

XX Purified nuclear factor of activated T cell protein - useful for

XX detecting protein expression and screening potential

XX immunosuppressants

XX Disclosure; Page 58-61; 100pp; English.

XX A cDNA clone, mNF-ATPQ1B1/A, of mouse preexisting nuclear factor

XX of activated T-cell protein (NF-ATp) was isolated from a T-cell

XX library. The open reading frame encoded the protein given in

XX AAR60252.

XX Sequence 890 AA;

Query Match 54.5%; Score 6; DB 15; Length 890;  
 Best Local Similarity 100.0%; Pred. No. 73;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OOOTAP 6  
 1111111  
 Db 695 qqlap 700

## RESULT 10

AAV26181  
 ID AAV26181 standard; Protein; 11 AA.

AC AAV26181;

PF 03-JUN-1993: 93DE-4318450.  
 XX  
 PR 11-JUN-1992: 92DE-4219111.  
 PR 25-NOV-1992: 92DE-4239567.  
 XX  
 PA (MERCK ) MERCK PATENT GMBH.  
 PI Schubert P, Neumann S, Pawelzik W, Linxweiler W, Burger C;  
 PI Hofmann G, Bubert A, Goebel W, Koehler S;  
 XX  
 DR WPI: 1993-406956/51.  
 XX  
 PT New primers for PCR detection of Listeria - including individual  
 PT species, also new peptide(s) for raising antibodies for  
 PT immunochemical detection  
 PS  
 XX Disclosure; Page 3; 19pp; German.  
 XX  
 CC The sequence is that of a Listeria p60 peptide epitope which  
 CC which may be used in the prodn. of antibodies for the detection  
 CC of Listeria by immunoassay (partic. ELISA). It may be used as  
 CC part of a method that allows determination of individual Listeria  
 CC species, esp. L. monocytogenes.  
 CC  
 SQ Sequence 9 AA:

Query Match 81.8%; Score 9; DB 14; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QQTAPKAPT 10  
 |||||  
 Db 1 qqtapkept 9

RESULT 6  
 AAR45162  
 ID AAR45162 standard; Protein: 23 AA.  
 XX  
 AC AAR45162;  
 XX  
 DT 16-JUN-1994 (first entry)  
 XX  
 DE Listeria p60 peptide epitope.  
 XX  
 KW Listeria monocytogenes; antibodies; immunoassay; conjugate.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..7  
 FT /note= "opt. 0-7 amino acid residues"  
 FT Region 17..23  
 FT /note= "opt. 0-7 amino acid residues"  
 XX  
 PN DE4318450-A.  
 XX  
 PD 16-DEC-1993.  
 XX  
 PF 03-JUN-1993: 93DE-4318450.  
 XX  
 PR 11-JUN-1992: 92DE-4219111.  
 PR 25-NOV-1992: 92DE-4239567.  
 XX  
 PA (MERCK ) MERCK PATENT GMBH.  
 PI Schubert P, Neumann S, Pawelzik W, Linxweiler W, Burger C;  
 PI Hofmann G, Bubert A, Goebel W, Koehler S;  
 XX  
 DR WPI: 1993-406956/51.  
 XX  
 PT New primers for PCR detection of Listeria - including individual

PT species, also new peptide(s) for raising antibodies for  
 PT immunochemical detection  
 XX  
 PS Claim 3; Page 4; 19pp; German.  
 XX  
 CC The sequence is that of a Listeria p60 peptide epitope which  
 CC which may be used in the prodn. of antibodies for the detection  
 CC of Listeria by immunoassay (partic. ELISA). It may be used as  
 CC part of a method that allows determination of individual Listeria  
 CC species, esp. L. monocytogenes.  
 CC  
 SQ Sequence 23 AA:

Query Match 81.8%; Score 9; DB 14; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 0.0023;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 QQTAPKAPT 10  
 |||||  
 Db 8 qqtapkept 16

RESULT 7  
 AAR73906  
 ID AAR73906 standard; peptide: 7 AA.  
 XX  
 AC AAR73906;  
 XX  
 DT 05-DEC-1995 (first entry)  
 XX  
 DE Listeria monocytogenes protein p60 precursor peptide 292-298.  
 XX  
 KW Listeria monocytogenes; protein p60 precursor; vaccine;  
 KW meningitis related homologous antigenic sequence; MRHAS; RV-1;  
 KW immunoassay; diagnosis; treatment; prophylactic; bacterial;  
 KW viral; peptide 292-298.  
 XX  
 OS Listeria monocytogenes.  
 XX  
 PN W09509232-A.  
 XX  
 PD 06-APR-1995.  
 XX  
 PF 28-SEP-1994: 94WO-CA00516.  
 XX  
 PR 28-SEP-1993: 93US-0127499.  
 XX  
 PA (SHAR//) SHARMA L R.  
 PA (VALS//) VAN ALSTYNE D.  
 XX  
 PI Sharma LR, Van Alstyne D;  
 XX  
 DR WPI: 1995-147431/19.  
 XX  
 PT New peptide(s) and corresp. antibodies for the treatment of  
 PT meningitis - the peptide(s) corresp. to homologous antigenic  
 PT sites on bacterial and viral agents and on chemokine(s), used for  
 PT detecting and preventing meningitis  
 PS  
 XX Claim 35; Page 77; 98pp; English.  
 CC AAR73913 is the Listeria monocytogenes protein p60 precursor. It  
 CC contains the meningitis related antigenic sequences (MRHAS) claimed  
 CC in AAR73891-R73894 and AAR73903-R73906, which are recognised by a  
 CC monoclonal antibody from the hybridoma Rubella virus (RV)-1. The  
 CC claimed MRHAS peptides may be used in immunoassays to diagnose the  
 CC presence of bacterial and/or viral meningitis agents in a sample,  
 CC or in prophylactic and therapeutic meningitis treatments. The  
 CC peptides may also be used as vaccines against meningitis.  
 CC  
 SQ Sequence 7 AA:



961	4	36.4	179	21	AA606292	Arabidopsis thalia
962	4	36.4	179	21	AA641095	zea mays protein f
963	4	36.4	179	21	AA58630	Protein regulating
964	4	36.4	179	22	AA846316	H. pylori HPS144 p
965	4	36.4	179	22	AA836630	Human FLEXIT-52 pr
966	4	36.4	179	22	AA836630	C-terminal portion
967	4	36.4	180	14	AA843883	Human clone 34 pro
968	4	36.4	180	17	AA843884	Human clone 34 pro
969	4	36.4	180	17	AA843884	Human clone 34 pro
970	4	36.4	180	17	AA843884	Human clone 34 pro
971	4	36.4	181	19	AA857883	Human GRP-binding
972	4	36.4	181	21	AA828367	Arabidopsis thalia
973	4	36.4	182	19	AA856799	S. thermophilus fl
974	4	36.4	182	20	AA858894	Neisseria meningit
975	4	36.4	183	10	AA833229	zea mays protein f
976	4	36.4	183	10	AA833229	Amino acid sequenc
977	4	36.4	183	11	AA806439	MG-4 antigen. AA
978	4	36.4	183	20	AA807984	Human secreted pro
979	4	36.4	183	21	AA806111	Arabidopsis thalia
980	4	36.4	183	21	AA819925	Arabidopsis thalia
981	4	36.4	184	19	AA856796	L. lactis fl portil
982	4	36.4	184	20	AA855124	C. pneumoniae prot
983	4	36.4	184	21	AA824729	Plant SPF encoded
984	4	36.4	184	21	AA816320	Eucalyptus grandis
985	4	36.4	184	21	AA807341	Arabidopsis thalia
986	4	36.4	184	21	AA813308	Arabidopsis thalia
987	4	36.4	184	21	AA839660	Arabidopsis thalia
988	4	36.4	184	21	AA850665	Arabidopsis thalia
989	4	36.4	185	21	AA842098	Arabidopsis thalia
990	4	36.4	185	21	AA858323	Arabidopsis thalia
991	4	36.4	186	21	AA850971	Arabidopsis thalia
992	4	36.4	186	21	AA850008	Neisseria gonorrhe
993	4	36.4	186	21	AA850008	Neisseria meningit
994	4	36.4	186	21	AA850010	Neisseria meningit
995	4	36.4	186	22	AA887419	Human gene 3 encod
996	4	36.4	187	18	AA820795	H. pylori outer me
997	4	36.4	187	21	AA833058	Pinus radiata tran
998	4	36.4	187	21	AA840385	Arabidopsis thalia
999	4	36.4	188	20	AA893946	Human regulatory m
1000	4	36.4	188	21	AA807997	Amino acid sequenc
					AA859574	Arabidopsis thalia

## ALIGNMENTS

RESULT 1

AA845171

ID AAR45171 standard; Protein: 11 AA.

AC AAR45171:

DT 16-JUN-1994 (first entry)

XX Listeria p60 peptide epitope.

DE Listeria monocytogenes; antibodies; immunosassay; conjugate.

XX

XX Synthetic.

OS DE4318450-A.

PN 16-DEC-1993.

PD 03-JUN-1993; 93DE-4318450.

PE 11-JUN-1992; 92DE-4219111.

PR 25-NOV-1992; 92DE-4239567.

XX (MERE ) MERCK PATENT GMBH.

PA Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;

PI Hofmann G, Bubert A, Goebel W, Koehler S;

XX

DR WPI: 1993-406956/51.

XX New primers for PCR detection of Listeria - including individual

PT species, also new peptide(s) for raising antibodies for

PT immunochemical detection

XX

PS Disclosure; Fig 2; 19pp; German.

XX

CC The sequence is that of a Listeria p60 peptide epitope which

CC which may be used in the prohn. of antibodies for the detection

CC of Listeria by Immunossay (partic. ELISA). It may be used as

CC part of a method that allows determination of individual Listeria

CC species, esp. L. monocytogenes.

XX

SQ Sequence 11 AA:

Query Match 100.0%; Score 11; DB 14; Length 11;

Best Local Similarity 100.0%; Pred. No. 1e-05;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OOOTAPKAPTE 11

DB 1 qgqtapkapte 11

RESULT 2

AA873894

ID AAR73894 standard; peptide: 21 AA.

AC AAR73894:

XX

DT 05-DEC-1995 (first entry)

DE Listeria monocytogenes protein P60 precursor peptide 285-305.

XX

KW Listeria monocytogenes; protein P60 precursor; vaccine;

KW meningitis related homologous antigenic sequence; MRHAS; RV-1;

KW immunossay; diagnosis; treatment; prophylactic; bacterial;

KW viral; peptide 285-305.

XX

OS Listeria monocytogenes.

XX

PN WO9509232-A.

PD 06-APR-1995.

XX

PE 28-SEP-1994; 94WO-CA00516.

XX

PR 28-SEP-1993; 93US-0127499.

XX

PA (SHAR/) SHARMA L R.

PA (VALS/) VAN ALSTYNE D.

PI Sharma LR, Van Alstyne D;

XX

DR WPI: 1995-147431/19.

XX

PT New peptide(s) and corresp. antibodies for the treatment of

PT meningitis - the peptide(s) corresp. to homologous antigenic

PT sites on bacterial and viral agents and on chemokine(s), used for

PT detecting and preventing meningitis

XX

PS Claim 34; Page 74; 98pp; English.

XX

CC AAR73913 is the Listeria monocytogenes protein P60 precursor. It

CC contains the meningitis related antigenic sequences (MRHAS) claimed

CC in AAR73891-873894 and AAR73903-873906, which are recognised by a

CC monoclonal antibody from the hybridoma Rubella virus (RV)-1. The

CC claimed MRHAS peptides may be used in immunossays to diagnose the

CC presence of bacterial and/or viral meningitis agents in a sample,

CC or in prophylactic and therapeutic meningitis treatments. The

CC peptides may also be used as vaccines against meningitis.

815	4	36.4	147	21	AAAG14003	Arabidopsis thalia
816	4	36.4	147	21	AAAG41096	zea mays protein f
817	4	36.4	147	21	AAAG45002	zea mays protein f
818	4	36.4	148	18	AAAM12693	Human cystatin E.
819	4	36.4	148	20	AAAY41588	Fragment of human
820	4	36.4	148	21	AAAB54088	Human pancreatic c
821	4	36.4	149	21	AAAG40950	zea mays protein f
822	4	36.4	149	21	AAAG49980	Arabidopsis thalia
823	4	36.4	149	21	AAAG54635	zea mays protein f
824	4	36.4	149	22	AAAT72673	Polyglutamine trac
825	4	36.4	150	20	AAAT28887	Lactobacillus brev
826	4	36.4	150	21	AAAG25512	Arabidopsis thalia
827	4	36.4	150	21	AAAG41597	Arabidopsis thalia
828	4	36.4	151	18	AAAM27857	Amino acid sequenc
829	4	36.4	151	18	AAAM25527	Staphylococcus aur
830	4	36.4	151	21	AAAG26405	Arabidopsis thalia
831	4	36.4	151	21	AAAG33230	zea mays protein f
832	4	36.4	151	21	AAAG33786	Arabidopsis thalia
833	4	36.4	151	21	AAAG35418	zea mays protein f
834	4	36.4	151	21	AAAG40734	zea mays protein f
835	4	36.4	152	12	AAAR12520	Adometase encoded
836	4	36.4	152	15	AAAR74824	Bacteriophage T3 S
837	4	36.4	152	16	AAAR75212	Bacteriophage T3 S
838	4	36.4	152	17	AAAR88606	S-adenosylmethioni
839	4	36.4	152	17	AAAR88611	S-adenosylmethioni
840	4	36.4	152	17	AAAR88601	S-adenosylmethioni
841	4	36.4	152	17	AAAR88603	Adometase (12V) SA
842	4	36.4	152	18	AAAM29147	S-adenosylmethioni
843	4	36.4	152	18	AAAM11594	Adometase. Bacter
844	4	36.4	152	20	AAAY48291	Human prostate can
845	4	36.4	152	20	AAAM30603	Tomato E4 promoter
846	4	36.4	152	20	AAAM30604	S-adenosylmethioni
847	4	36.4	152	21	AAAG44025	Arabidopsis thalia
848	4	36.4	152	21	AAAG54637	zea mays protein f
849	4	36.4	153	20	AAAY74188	Human prostate tum
850	4	36.4	153	21	AAAY53589	Human preproinsuli
851	4	36.4	154	21	AAAG40970	zea mays protein f
852	4	36.4	154	21	AAAG44024	Arabidopsis thalia
853	4	36.4	154	22	AAAB30324	Helicobacter pylori
854	4	36.4	155	19	AAAM72382	Arabidopsis pathog
855	4	36.4	155	22	AAAB42960	S. cerevisiae YFR0
856	4	36.4	156	20	AAAY60060	Human endometrium
857	4	36.4	156	20	AAAM90885	Non-B, non-C, non-
858	4	36.4	156	21	AAAG18865	zea mays protein f
859	4	36.4	156	21	AAAG23571	zea mays protein f
860	4	36.4	156	21	AAAG54607	Human breast cance
861	4	36.4	156	22	AAAB63350	Mouse CIP1 protein
862	4	36.4	157	19	AAAM58594	Maize CIP1 protein
863	4	36.4	157	21	AAAB10276	Marine adult thymu
864	4	36.4	157	21	AAAB10276	Arabidopsis thalia
865	4	36.4	157	21	AAAG11620	zea mays protein f
866	4	36.4	157	21	AAAG27172	Yeast NUP2 protein
867	4	36.4	158	20	AAAY08804	Yeast NUP2 protein
868	4	36.4	158	20	AAAY08832	Arabidopsis thalia
869	4	36.4	158	21	AAAB33212	Pinus radiata tran
870	4	36.4	158	21	AAAG04324	Arabidopsis thalia
871	4	36.4	158	21	AAAG45297	Arabidopsis thalia
872	4	36.4	159	20	AAAY00318	Human secreted pro
873	4	36.4	159	21	AAAG08936	Arabidopsis thalia
874	4	36.4	159	21	AAAG10210	Arabidopsis thalia
875	4	36.4	159	21	AAAG11619	Arabidopsis thalia
876	4	36.4	159	21	AAAG40733	zea mays protein f
877	4	36.4	159	21	AAAG48761	Arabidopsis thalia
878	4	36.4	159	21	AAAG53229	Arabidopsis thalia
879	4	36.4	160	15	AAAR60319	Pre-pro-vasoactive
880	4	36.4	160	15	AAAR60320	Human colon cancer
881	4	36.4	160	21	AAAB53652	Human A-proinsul
882	4	36.4	160	21	AAAB15419	Arabidopsis thalia
883	4	36.4	160	21	AAAG08107	Arabidopsis thalia
884	4	36.4	160	21	AAAG19894	Arabidopsis thalia
885	4	36.4	160	21	AAAG43222	Arabidopsis thalia
886	4	36.4	160	21	AAAG54875	Arabidopsis thalia
887	4	36.4	161	21	AAAB54206	Human pancreatic c
888	4	36.4	161	21	AAAG04244	Arabidopsis thalia
889	4	36.4	161	21	AAAG06112	Arabidopsis thalia
890	4	36.4	161	21	AAAG45945	Bacillus stearothe
891	4	36.4	162	20	AAAY50046	Human RNA-associat
892	4	36.4	162	21	AAAY70241	H. pylori ORF 14c
893	4	36.4	163	18	AAAM55291	H. Influenzae Yigu
894	4	36.4	163	20	AAAY41231	Human cancer assoc
895	4	36.4	163	21	AAAB43807	Arabidopsis thalia
896	4	36.4	163	21	AAAG08106	Protein G variant.
897	4	36.4	164	10	AAAP94787	Protein G variant.
898	4	36.4	164	11	AAAR07011	Type 9 Gx7809 prot
899	4	36.4	164	12	AAAR10015	IgG-binding Strept
900	4	36.4	164	15	AAAS53299	Plant SPD encoded
901	4	36.4	164	21	AAAB24872	Arabidopsis thalia
902	4	36.4	165	21	AAAG11618	Arabidopsis thalia
903	4	36.4	165	21	AAAG15310	Arabidopsis thalia
904	4	36.4	165	21	AAAG34537	zea mays protein f
905	4	36.4	165	21	AAAB57273	Human gene 72-enco
906	4	36.4	166	21	AAAB58526	Human prostate can
907	4	36.4	166	21	AAAG07157	Arabidopsis thalia
908	4	36.4	166	21	AAAG33328	zea mays protein f
909	4	36.4	166	21	AAAG48809	Arabidopsis thalia
910	4	36.4	167	21	AAAG38554	Arabidopsis thalia
911	4	36.4	168	20	AAAY50050	Mycobacterium tube
912	4	36.4	168	21	AAAB24730	Plant SPD encoded
913	4	36.4	168	21	AAAG29971	Arabidopsis thalia
914	4	36.4	169	21	AAAG08001	Arabidopsis thalia
915	4	36.4	169	21	AAAG19850	Arabidopsis thalia
916	4	36.4	169	21	AAAG19926	Arabidopsis thalia
917	4	36.4	169	21	AAAG55909	Arabidopsis thalia
918	4	36.4	169	21	AAAG60711	Arabidopsis thalia
919	4	36.4	170	21	AAAY74833	Neisseria gonorrhe
920	4	36.4	171	21	AAAB42103	Human ORFX ORF1867
921	4	36.4	171	21	AAAY95649	Neisseria meningit
922	4	36.4	171	21	AAAG06293	Arabidopsis thalia
923	4	36.4	171	21	AAAG39723	Arabidopsis thalia
924	4	36.4	171	21	AAAG60024	Arabidopsis thalia
925	4	36.4	172	21	AAAY99567	Vibrio cholerae lu
926	4	36.4	172	21	AAAG06271	Arabidopsis thalia
927	4	36.4	172	21	AAAG23200	Arabidopsis thalia
928	4	36.4	172	21	AAAG26404	Arabidopsis thalia
929	4	36.4	172	21	AAAG50704	Arabidopsis thalia
930	4	36.4	172	21	AAAG58325	Arabidopsis thalia
931	4	36.4	172	21	AAAG61127	Arabidopsis thalia
932	4	36.4	173	21	AAAG06270	Arabidopsis thalia
933	4	36.4	173	21	AAAG58324	Arabidopsis thalia
934	4	36.4	173	21	AAAG61126	Arabidopsis thalia
935	4	36.4	174	21	AAAG10209	Arabidopsis thalia
936	4	36.4	174	21	AAAG33901	Arabidopsis thalia
937	4	36.4	174	21	AAAG48760	Arabidopsis thalia
938	4	36.4	174	21	AAAG58553	Arabidopsis thalia
939	4	36.4	174	22	AAAB79407	Corynebacterium gl
940	4	36.4	175	21	AAAY81624	Streptococcus pneu
941	4	36.4	175	21	AAAB57224	Human prostate can
942	4	36.4	176	21	AAAB58426	Lung cancer associ
943	4	36.4	176	21	AAAG20227	Arabidopsis thalia
944	4	36.4	176	21	AAAG33900	Arabidopsis thalia
945	4	36.4	176	21	AAAG41596	Arabidopsis thalia
946	4	36.4	177	9	AAAP81164	Sequence of protei
947	4	36.4	177	20	AAAY33001	Human serine prote
948	4	36.4	177	21	AAAG32029	Arabidopsis thalia
949	4	36.4	178	13	AAAR20231	Human endothelin-2
950	4	36.4	178	13	AAAR23784	Precursor ER-2 seq
951	4	36.4	178	20	AAAY58888	Human normal uteru
952	4	36.4	178	20	AAAY04279	Borrelia afzelii A
953	4	36.4	178	21	AAAG41049	zea mays protein f
954	4	36.4	178	21	AAAG45944	Arabidopsis thalia
955	4	36.4	179	18	AAAM20394	H. pylori outer me
956	4	36.4	179	18	AAAM24651	H. pylori outer me
957	4	36.4	179	20	AAAM89813	Protein encoded by
958	4	36.4	179	20	AAAM89813	Expressed antigen
959	4	36.4	179	21	AAAB43346	Human ORFX ORF3110
960	4	36.4	179	21	AAAG04243	Arabidopsis thalia



669	4	36.4	120	21	AA643216	Arabidopsis thalia	742	4	36.4	133	10	AA694786	Protein G variant.
670	4	36.4	120	21	AA652566	Arabidopsis thalia	743	4	36.4	133	11	AA607005	Protein G variant.
671	4	36.4	120	21	AA603428	Human secreted pro	744	4	36.4	133	12	AA610010	Type 5 GX7809 prot
672	4	36.4	121	21	AA640960	Human ORFX ORF724	745	4	36.4	133	14	AA634012	dTDP-4'-keto-6-deox
673	4	36.4	121	21	AA619852	Arabidopsis thalia	746	4	36.4	133	15	AA653286	IgG-binding Strept
674	4	36.4	121	21	AA654575	Zea mays protein f	747	4	36.4	133	20	AA601132	Wilo-type Kexatrin
675	4	36.4	122	16	AA681442	Hepatitis GB virus	748	4	36.4	133	21	AA620228	Arabidopsis thalia
676	4	36.4	122	20	AA625862	Human secreted pro	749	4	36.4	133	21	AA627173	Zea mays protein f
677	4	36.4	122	21	AA623551	Invertebrate/thiored	750	4	36.4	133	21	AA641136	Zea mays protein f
678	4	36.4	122	21	AA618866	Zea mays protein f	751	4	36.4	133	21	AA648755	Arabidopsis thalia
679	4	36.4	122	21	AA654608	Zea mays protein f	752	4	36.4	134	11	AA604087	Protein encoded by
680	4	36.4	122	21	AA654636	Zea mays protein f	753	4	36.4	134	11	AA606603	Human Brain Natriu
681	4	36.4	122	21	AA608928	Human secreted pro	754	4	36.4	134	15	AA656036	Mutant human Inter
682	4	36.4	122	21	AA609019	Hepatitis GB virus	755	4	36.4	134	20	AA605325	Human gamma-BNP pr
683	4	36.4	123	15	AA654984	Iron oxide binding	756	4	36.4	134	21	AA616614	Bacteriophage 192
684	4	36.4	123	21	AA612977	Arabidopsis thalia	757	4	36.4	134	21	AA632664	Zea mays protein f
685	4	36.4	123	21	AA633788	Arabidopsis thalia	758	4	36.4	134	21	AA657449	Arabidopsis thalia
686	4	36.4	123	21	AA654874	Arabidopsis thalia	759	4	36.4	134	21	AA657449	Arabidopsis thalia
687	4	36.4	123	21	AA657374	Drosophila Shaker	760	4	36.4	135	20	AA648498	Human BNP preprope
688	4	36.4	123	21	AA650565	Drosophila Shaker	761	4	36.4	135	21	AA643215	Human breast tumou
689	4	36.4	124	20	AA66182	Human bladder tumo	762	4	36.4	136	16	AA664971	Arabidopsis thalia
690	4	36.4	125	20	AA629529	Human lung tumour	763	4	36.4	136	20	AA637091	Gag region p19 pro
691	4	36.4	125	21	AA657166	Human prostate can	764	4	36.4	136	21	AA658506	Amino acid sequenc
692	4	36.4	125	21	AA644425	Human lung tumour	765	4	36.4	136	21	AA658506	Lung cancer associ
693	4	36.4	125	21	AA611722	Arabidopsis thalia	766	4	36.4	136	21	AA622380	Arabidopsis thalia
694	4	36.4	125	21	AA633329	Arabidopsis thalia	767	4	36.4	136	21	AA624561	Arabidopsis thalia
695	4	36.4	125	21	AA644972	Zea mays protein f	768	4	36.4	136	21	AA638985	Arabidopsis thalia
696	4	36.4	125	21	AA667284	Zea mays protein f	769	4	36.4	136	21	AA641099	Arabidopsis thalia
697	4	36.4	126	14	AA643681	Signature pattern	770	4	36.4	136	21	AA601430	Zea mays protein f
698	4	36.4	126	14	AA643682	PBI.3/Humanised 11	771	4	36.4	136	21	AA601430	Human secreted pro
699	4	36.4	126	14	AA643682	PBI.3/Humanised 11	772	4	36.4	137	14	AA642796	Partial B4 Ab vari
700	4	36.4	126	15	AA62677	CY1748RLD VL regio	773	4	36.4	137	20	AA659678	Secreted protein 1
701	4	36.4	126	15	AA62677	CY1748RLD VL regio	774	4	36.4	137	21	AA654558	Zea mays protein f
702	4	36.4	126	15	AA62677	CY1748RLD VL regio	775	4	36.4	137	21	AA646506	A human small prol
703	4	36.4	126	18	AA628201	CY1748RLC VL regio	776	4	36.4	137	22	AA631675	Amino acid sequenc
704	4	36.4	126	18	AA628201	Staphylococcus aur	777	4	36.4	138	20	AA612378	Human 5' EST sece
705	4	36.4	126	21	AA608937	Arabidopsis thalia	778	4	36.4	138	20	AA641359	Human 5' EST sece
706	4	36.4	126	21	AA642059	Arabidopsis thalia	779	4	36.4	138	21	AA619895	Arabidopsis thalia
707	4	36.4	127	21	AA628907	Arabidopsis thalia	780	4	36.4	138	21	AA643223	Arabidopsis thalia
708	4	36.4	127	21	AA656716	Amino acid sequenc	781	4	36.4	138	21	AA653949	Arabidopsis thalia
709	4	36.4	128	20	AA629530	Human lung tumour	782	4	36.4	138	20	AA692366	MAS-like G protein
710	4	36.4	128	21	AA654340	Human pancreatic c	783	4	36.4	139	20	AA652779	Chlamydia pneumoni
711	4	36.4	128	21	AA644427	Human lung tumour	784	4	36.4	139	21	AA642428	Human ORFX ORF2192
712	4	36.4	128	21	AA619851	Arabidopsis thalia	785	4	36.4	139	21	AA633787	Arabidopsis thalia
713	4	36.4	128	21	AA656527	Human gene 72-enco	786	4	36.4	139	21	AA641154	Zea mays protein f
714	4	36.4	128	21	AA653382	Human 5' EST relat	787	4	36.4	139	21	AA647096	Arabidopsis thalia
715	4	36.4	128	22	AA648027	Light chain sequen	788	4	36.4	139	22	AA654522	Zea mays protein f
716	4	36.4	129	10	AA694792	Protein G variant.	789	4	36.4	139	22	AA679482	Corynebacterium gl
717	4	36.4	129	12	AA610016	Type 11 GX7809 pro	790	4	36.4	140	20	AA659723	Secreted protein 6
718	4	36.4	129	15	AA654716	IgG-binding Strept	791	4	36.4	140	20	AA65347	Human adult testis
719	4	36.4	129	21	AA640968	Zea mays protein f	792	4	36.4	140	21	AA618734	Zea mays protein f
720	4	36.4	130	20	AA697713	Staphylococcus aur	793	4	36.4	140	21	AA659789	MMpP-MMP20-(His
721	4	36.4	130	21	AA640588	Human ORFX ORF352	794	4	36.4	141	22	AA668882	Human RECAP polype
722	4	36.4	130	21	AA644026	Arabidopsis thalia	795	4	36.4	141	14	AA642797	Human RECAP polype
723	4	36.4	130	21	AA654638	Arabidopsis thalia	796	4	36.4	141	17	AA64413	Partial B13/B14 Ab
724	4	36.4	130	22	AA656733	Zea mays protein f	797	4	36.4	141	17	AA64413	Met-KGF delta-N23/
725	4	36.4	130	22	AA679798	Amino acid sequenc	798	4	36.4	141	19	AA614837	Kerittinocyte growt
726	4	36.4	130	22	AA680017	Corynebacterium gl	799	4	36.4	141	20	AA65726	Kerittinocyte growt
727	4	36.4	130	22	AA646816	C. glutamicum folb	800	4	36.4	142	21	AA625405	Amino acid sequenc
728	4	36.4	131	22	AA648028	Light chain sequen	801	4	36.4	142	21	AA615311	Pinus radiata cell
729	4	36.4	131	10	AA694791	Protein G variant.	802	4	36.4	142	21	AA65751	Arabidopsis thalia
730	4	36.4	131	12	AA607008	Protein G variant.	803	4	36.4	143	18	AA656751	Arabidopsis thalia
731	4	36.4	131	15	AA610017	Type 10 GX7809 pro	804	4	36.4	143	21	AA608108	S. pneumoniae enol
732	4	36.4	131	15	AA653300	IgG-binding Strept	805	4	36.4	144	21	AA620404	Arabidopsis thalia
733	4	36.4	131	15	AA648625	Sequence of mouse	806	4	36.4	144	21	AA641050	Arabidopsis thalia
734	4	36.4	131	18	AA619533	HSV 863 antibody 1	807	4	36.4	144	22	AA680201	Zea mays protein f
735	4	36.4	131	18	AA626652	64-863 antibody HS	808	4	36.4	145	20	AA629459	Corynebacterium gl
736	4	36.4	131	20	AA637836	Amino acid sequenc	809	4	36.4	145	20	AA644433	Human lung tumour
737	4	36.4	131	20	AA658555	Arabidopsis thalia	810	4	36.4	145	21	AA620403	Human lung tumour
738	4	36.4	132	20	AA607463	Anti-HIV-1 gp120 a	811	4	36.4	146	18	AA635500	Arabidopsis thalia
739	4	36.4	132	21	AA612012	Arabidopsis thalia	812	4	36.4	146	21	AA616312	Streptomyces venez
740	4	36.4	132	21	AA615312	Arabidopsis thalia	813	4	36.4	146	22	AA657448	Arabidopsis thalia
741	4	36.4	132	21	AA694946	Human secreted pro	814	4	36.4	147	19	AA688537	Haemophilus influe
					AA658203	Canine Interleukin						AA672404	Arabidopsis thalia

523	4	36.4	107	21	AAG41155	Zea mays protein f	596	4	36.4	112	21	AAB39522	Anti-IL12 antibody
524	4	36.4	107	21	AAV78860	Streptomycin subti	597	4	36.4	112	21	AAB39524	Anti-IL12 antibody
525	4	36.4	107	21	AAV56680	Anti-erythropoietin	598	4	36.4	112	21	AAB39526	Anti-IL12 antibody
526	4	36.4	108	21	AAB19749	Erythropoietin rec	599	4	36.4	112	21	AAB39528	Anti-IL12 antibody
527	4	36.4	108	21	AAB19750	Erythropoietin rec	600	4	36.4	112	21	AAB39530	Anti-IL12 antibody
528	4	36.4	108	21	AAB19751	Erythropoietin rec	601	4	36.4	112	21	AAB39532	Anti-IL12 antibody
529	4	36.4	108	21	AAB19752	Erythropoietin rec	602	4	36.4	112	21	AAB39534	Anti-IL12 antibody
530	4	36.4	108	21	AAG22382	Arabidopsis thalia	603	4	36.4	112	21	AAB39536	Anti-IL12 antibody
531	4	36.4	108	21	AAG02229	Human secreted pro	604	4	36.4	112	21	AAB39538	Anti-IL12 antibody
532	4	36.4	108	21	AAV58205	Canine mature inte	605	4	36.4	112	21	AAB39540	Anti-IL12 antibody
533	4	36.4	109	11	AAV07009	Protein G variant.	606	4	36.4	112	21	AAB39542	Anti-IL12 antibody
534	4	36.4	109	12	AAV15646	Thioredoxin. Salm	607	4	36.4	112	21	AAB39544	Anti-IL12 antibody
535	4	36.4	109	16	AAV75767	Hls-patch thioredo	608	4	36.4	112	21	AAB39546	Anti-IL12 antibody
536	4	36.4	109	16	AAV75768	Histidine patch-2	609	4	36.4	112	21	AAB39548	Anti-IL12 antibody
537	4	36.4	109	16	AAV75769	E. coli thioredoxi	610	4	36.4	112	21	AAB39550	Anti-IL12 antibody
538	4	36.4	109	18	AAV37348	Immunoglobulin C-1	611	4	36.4	112	21	AAB39552	Anti-IL12 antibody
539	4	36.4	109	18	AAV27547	Human Ab 11ght cha	612	4	36.4	112	21	AAB39554	Anti-IL12 antibody
540	4	36.4	109	20	AAV25383	Plasmd d PRX thior	613	4	36.4	112	21	AAB39556	Anti-IL12 antibody
541	4	36.4	109	20	AAV05567	HIV-1 group O isol	614	4	36.4	112	21	AAB39558	Anti-IL12 antibody
542	4	36.4	109	20	AAV02639	Thioredoxin protei	615	4	36.4	112	21	AAB39560	Anti-IL12 antibody
543	4	36.4	109	20	AAV88817	Polypeptide fragme	616	4	36.4	112	21	AAB39562	Anti-IL12 antibody
544	4	36.4	109	21	AAV56903	Human prostate can	617	4	36.4	112	21	AAV08002	Pinus radiata tran
545	4	36.4	109	21	AAV23553	Thioredoxin protei	618	4	36.4	112	21	AAV94933	Arabidopsis thalia
546	4	36.4	109	21	AAV24873	Plant SDF encoded	619	4	36.4	112	22	AAV48019	Human secreted pro
547	4	36.4	109	21	AAV19832	Arabidopsis thalia	620	4	36.4	113	14	AAV34276	Glamda-1 light ch
548	4	36.4	109	21	AAV32030	Arabidopsis thalia	621	4	36.4	113	14	AAV34277	Human Vlambdal-1-1
549	4	36.4	109	21	AAV32665	Zea mays protein f	622	4	36.4	113	14	AAV34278	Human Vlambdal-1-1
550	4	36.4	109	21	AAV45707	Arabidopsis thalia	623	4	36.4	113	19	AAV75885	Peptide inhibitor
551	4	36.4	109	21	AAV00954	Human secreted pro	624	4	36.4	113	19	AAV75885	Zea mays protein f
552	4	36.4	109	21	AAV95204	Anti-platelet glyc	625	4	36.4	113	21	AAV18475	Human secreted pro
553	4	36.4	109	21	AAV95205	Anti-platelet glyc	626	4	36.4	113	21	AAV78685	Streptomycin subti
554	4	36.4	109	22	AAV80367	Human prostate can	627	4	36.4	113	22	AAV63888	Human prostate can
555	4	36.4	110	14	AAV33331	Variant IGE - muta	628	4	36.4	114	20	AAV34698	Chlamydia pneumoni
556	4	36.4	110	16	AAV80621	Human tumor-associ	629	4	36.4	114	21	AAV07906	Arabidopsis thalia
557	4	36.4	110	19	AAV41425	Key2 protease inhi	630	4	36.4	114	21	AAV54523	Zea mays protein f
558	4	36.4	110	20	AAV01128	Wild-type Kexlatin	631	4	36.4	115	20	AAV41019	RKSI protein RGS r
559	4	36.4	110	21	AAV40424	Human ORFX ORF18	632	4	36.4	115	21	AAV45003	Zea mays protein f
560	4	36.4	110	21	AAV15654	Arabidopsis thalia	633	4	36.4	117	14	AAV33272	T cell receptor al
561	4	36.4	110	21	AAV20405	Arabidopsis thalia	634	4	36.4	117	14	AAV33273	T cell receptor al
562	4	36.4	110	21	AAV40977	Zea mays protein f	635	4	36.4	117	14	AAV33274	T cell receptor al
563	4	36.4	110	21	AAV95182	Anti-platelet glyc	636	4	36.4	117	14	AAV33275	T cell receptor al
564	4	36.4	111	12	AAV12263	Anti-human RND FOG	637	4	36.4	117	21	AAV01398	Neuron-associated
565	4	36.4	111	17	AAV03175	START2 DNA binding	638	4	36.4	117	21	AAV14004	Arabidopsis thalia
566	4	36.4	111	17	AAV88506	Vilamda for antib	639	4	36.4	117	21	AAV15989	Arabidopsis thalia
567	4	36.4	111	17	AAV89148	Human anti-Pseudom	640	4	36.4	117	21	AAV40973	Zea mays protein f
568	4	36.4	111	17	AAV92795	Murine p600. Mus	641	4	36.4	117	21	AAV03408	Human secreted pro
569	4	36.4	111	18	AAV13528	Anti-melanoma 11gh	642	4	36.4	117	21	AAV03857	Human secreted pro
570	4	36.4	111	18	AAV13515	Anti-melanoma 11gh	643	4	36.4	117	21	AAV03859	Human secreted pro
571	4	36.4	111	18	AAV13513	Anti-melanoma 11gh	644	4	36.4	118	20	AAV12136	Human 5' EST secre
572	4	36.4	111	18	AAV08489	C6 human sfv antib	645	4	36.4	118	21	AAV18933	Amino acid sequenc
573	4	36.4	111	18	AAV08491	C6 human sfv antib	646	4	36.4	118	21	AAV18935	Beta megaspermia p
574	4	36.4	111	18	AAV08492	C6 human sfv antib	647	4	36.4	118	21	AAV08050	Arabidopsis thalia
575	4	36.4	111	19	AAV52223	Antibody LD2-1-VL	648	4	36.4	118	21	AAV22381	Arabidopsis thalia
576	4	36.4	111	20	AAV90281	Human anti-GPIIb/I	649	4	36.4	118	21	AAV30057	Arabidopsis thalia
577	4	36.4	111	20	AAV90279	Human anti-GPIIb/I	650	4	36.4	119	9	AAV81366	Light chain variab
578	4	36.4	111	20	AAV01129	Kexlatin I mutant	651	4	36.4	119	21	AAV11074	Arabidopsis thalia
579	4	36.4	111	20	AAV01130	Kexlatin I mutant	652	4	36.4	119	21	AAV19518	Arabidopsis thalia
580	4	36.4	111	20	AAV01131	Kexlatin I mutant	653	4	36.4	119	21	AAV25089	Arabidopsis thalia
581	4	36.4	111	21	AAV40700	Human ORFX ORF464	654	4	36.4	119	21	AAV25513	Arabidopsis thalia
582	4	36.4	111	22	AAV48025	Glamda-1A single	655	4	36.4	119	21	AAV26139	Zea mays protein f
583	4	36.4	112	16	AAV76937	PHF-tau (143-254)	656	4	36.4	119	21	AAV34376	Arabidopsis thalia
584	4	36.4	112	18	AAV13523	Anti-melanoma 11gh	657	4	36.4	119	21	AAV35119	Arabidopsis thalia
585	4	36.4	112	18	AAV08488	C6 human sfv antib	658	4	36.4	119	21	AAV38181	Arabidopsis thalia
586	4	36.4	112	19	AAV06475	Human lung cancer	659	4	36.4	119	21	AAV38472	Arabidopsis thalia
587	4	36.4	112	19	AAV40534	Antibody HB4C5 119	660	4	36.4	119	21	AAV40975	Zea mays protein f
588	4	36.4	112	20	AAV60037	Human endometrium	661	4	36.4	119	21	AAV40987	Zea mays protein f
589	4	36.4	112	21	AAV52208	Human anti-HBs ant	662	4	36.4	119	21	AAV49989	Arabidopsis thalia
590	4	36.4	112	21	AAV39492	Anti-IL12 antibo	663	4	36.4	119	21	AAV15922	Arabidopsis thalia
591	4	36.4	112	21	AAV39500	Anti-IL12 antibo	664	4	36.4	119	22	AAV78992	Anti-IL12 antibo
592	4	36.4	112	21	AAV39508	Anti-IL12 antibo	665	4	36.4	120	15	AAV47044	C. glutamicum SRT
593	4	36.4	112	21	AAV39516	Anti-IL12 antibo	666	4	36.4	120	20	AAV48601	Liporotein. Synt
594	4	36.4	112	21	AAV39518	Anti-IL12 antibo	667	4	36.4	120	21	AAV24563	Human breast tumo
595	4	36.4	112	21	AAV39520	Anti-IL12 antibo	668	4	36.4	120	21	AAV28908	Arabidopsis thalia

377	4	36.4	70	21	AAB52587	Helicobacter pylori	450	4	36.4	96	21	AA626874	Zea mays protein f
378	4	36.4	71	18	AAW20467	H. pylori cytoplas	451	4	36.4	96	21	AA654577	Zea mays protein f
379	4	36.4	71	18	AAW24669	H. pylori cytoplas	452	4	36.4	97	21	AA604325	Arabidopsis thalia
380	4	36.4	71	20	AAV41589	Fragment of human	453	4	36.4	97	21	AA623454	Arabidopsis thalia
381	4	36.4	71	21	AA602695	Human secreted pro	454	4	36.4	97	21	AA645298	Arabidopsis thalia
382	4	36.4	72	21	AAB34764	Human secreted pro	455	4	36.4	98	14	AA634275	Human TNF binding
383	4	36.4	72	21	AA626400	Arabidopsis thalia	456	4	36.4	98	18	AAW14540	Streptococcus pneu
384	4	36.4	72	21	AA638456	Arabidopsis thalia	457	4	36.4	98	21	AAAB18934	Amino acid sequenc
385	4	36.4	73	21	AA607980	Arabidopsis thalia	458	4	36.4	98	21	AAAB40142	Anti-hi112 antibod
386	4	36.4	73	21	AA608477	Arabidopsis thalia	459	4	36.4	98	21	AAAB40143	Anti-hi112 antibod
387	4	36.4	73	22	AAB48747	Mouse liver growth	460	4	36.4	98	21	AAAB40144	Anti-hi112 antibod
388	4	36.4	74	18	AAW20946	H. pylori cytoplas	461	4	36.4	98	21	AAAB40145	Anti-hi112 antibod
389	4	36.4	74	20	AAV33812	Scsme-YOR159C a ye	462	4	36.4	98	21	AAAB40149	Anti-hi112 antibod
390	4	36.4	74	21	AA619051	Zea mays protein f	463	4	36.4	98	21	AAAG15354	Arabidopsis thalia
391	4	36.4	74	21	AA626399	Arabidopsis thalia	464	4	36.4	98	21	AAAG34149	Zea mays protein f
392	4	36.4	74	21	AA603601	Human secreted pro	465	4	36.4	98	21	AAAG01380	Human secreted pro
393	4	36.4	75	15	AAW42018	Clone BL341.4 prot	466	4	36.4	99	21	AAAB40147	Anti-hi112 antibod
394	4	36.4	75	20	AAV08625	Human secreted pro	467	4	36.4	99	21	AAAB40148	Anti-hi112 antibod
395	4	36.4	75	20	AAV11795	Human 5' EST secre	468	4	36.4	99	21	AAAG41137	Zea mays protein f
396	4	36.4	75	20	AAV01332	Modified sp6 RNA p	469	4	36.4	99	21	AAAG01078	Human secreted pro
397	4	36.4	75	21	AAV01363	Modified sp6 RNA p	470	4	36.4	99	21	AAV76755	Human protein kina
398	4	36.4	75	21	AAAB1141	Human ORFX ORP905	471	4	36.4	100	21	AAAB40576	Human ORFX ORP340
399	4	36.4	75	21	AAV67316	Human secreted pro	472	4	36.4	100	21	AAV93883	Amino acid sequenc
400	4	36.4	76	21	AAAB1661	Human ORFX ORP1425	473	4	36.4	101	14	AAAB38604	NEWM light chain.
401	4	36.4	76	21	AA654780	Arabidopsis thalia	474	4	36.4	101	20	AAV12102	Human 5' EST secre
402	4	36.4	76	21	AA658573	Arabidopsis thalia	475	4	36.4	102	21	AAAB41432	Human ORFX ORP196
403	4	36.4	76	21	AA603551	Human secreted pro	476	4	36.4	102	21	AA608003	Arabidopsis thalia
404	4	36.4	77	11	AA607049	Alkaline phosphata	477	4	36.4	102	21	AA622573	Zea mays protein f
405	4	36.4	77	20	AAV60212	Human endometrium	478	4	36.4	102	21	AA603918	Human secreted pro
406	4	36.4	77	20	AAV11964	Human 5' EST secre	479	4	36.4	102	21	AAV56683	Humanized 369 Ligh
407	4	36.4	77	21	AA638455	Arabidopsis thalia	480	4	36.4	102	22	AAAB63506	Human gastric canc
408	4	36.4	78	21	AA628397	Ubiquitin-like dom	481	4	36.4	103	14	AAAB38605	NEWM light chain.
409	4	36.4	79	20	AAW67979	Arabidopsis thalia	482	4	36.4	103	15	AAAE2037	Light chain variab
410	4	36.4	79	21	AA607979	Arabidopsis thalia	483	4	36.4	103	18	AAAW21635	NEWM human antibod
411	4	36.4	80	21	AA658711	Arabidopsis thalia	484	4	36.4	103	18	AAAB32442	Mycobacterium tube
412	4	36.4	80	21	AAB54219	Human pancreatic c	485	4	36.4	103	18	AAW32374	Mycobacterium tube
413	4	36.4	80	21	AA661350	Arabidopsis thalia	486	4	36.4	103	19	AAW81677	M. tuberculosis im
414	4	36.4	81	9	AAAP81320	MMP signal and N-t	487	4	36.4	103	19	AAW64314	Mycobacterium tube
415	4	36.4	81	20	AAW67980	Ubiquitin-like dom	488	4	36.4	103	20	AAW58484	Human NEWM antibod
416	4	36.4	81	21	AAAG11420	Arabidopsis thalia	489	4	36.4	103	20	AAV60148	Human endometrium
417	4	36.4	82	15	AAAE58336	Hypotensive polype	490	4	36.4	103	20	AAV39116	M. tuberculosis an
418	4	36.4	83	11	AAAR05708	Pseudomonas aerugi	491	4	36.4	103	20	AAV38979	M. tuberculosis re
419	4	36.4	83	16	AAAR6924	Pseudomonas aerugi	492	4	36.4	103	21	AAAB41048	Human ORFX ORF812
420	4	36.4	83	16	AAW41427	Kex2 protease inh	493	4	36.4	103	21	AAAG18812	Zea mays protein f
421	4	36.4	83	20	AAV73979	Human prostate tum	494	4	36.4	103	21	AAAC59656	Arabidopsis thalia
422	4	36.4	83	21	AA656133	Arabidopsis thalia	495	4	36.4	104	16	AAAR80087	Human derived ligh
423	4	36.4	84	8	AAAP70614	Sequence encoded b	496	4	36.4	104	20	AAW95485	Human-derived RT3
424	4	36.4	84	20	AAV07794	Human secreted pro	497	4	36.4	104	21	AAAG33330	Zea mays protein f
425	4	36.4	85	9	AAAP80331	Sequence of polype	498	4	36.4	105	9	AAAP81776	Sequence encoded b
426	4	36.4	85	17	AAAR9425	Mucin-derived prot	499	4	36.4	105	11	AAAR04028	R gene product of
427	4	36.4	85	21	AA602578	Human secreted pro	500	4	36.4	105	13	AAAR20601	ROD HIV-2 R protei
428	4	36.4	85	21	AAV69790	MMPsp-MMPm20-(His	501	4	36.4	105	13	AAAR20603	ROD HIV-2 Y protei
429	4	36.4	87	20	AAV72569	Human secreted pro	502	4	36.4	105	17	AAAR94545	Native Vpr protein
430	4	36.4	88	21	AAAB1338	Human ORFX ORF102	503	4	36.4	105	17	AAAR89427	Mucin-derived prot
431	4	36.4	88	21	AAAB7257	Soybean cyclin-dep	504	4	36.4	105	19	AAW40579	Antibody LP2-11-VL
432	4	36.4	88	21	AAV56651	Partial peptide fr	505	4	36.4	105	19	AAW52231	Antibody LP2-17-VL
433	4	36.4	89	21	AAV56668	Partial peptide fr	506	4	36.4	105	19	AAW52235	Lactobacillus brev
434	4	36.4	89	16	AAAR71124	SG-2. Synthetic.	507	4	36.4	105	20	AAV28886	Human lambda-CL do
435	4	36.4	89	20	AAV73958	Human prostate tum	508	4	36.4	105	20	AAV08746	Human lambda-CL do
436	4	36.4	90	20	AAV35733	Chlamydia pneumoni	509	4	36.4	105	20	AAW92426	Human lambda prote
437	4	36.4	90	21	AAAB58188	Lung cancer associ	510	4	36.4	105	21	AAAB4264	Human ORFX ORF2408
438	4	36.4	90	21	AAAG28396	Arabidopsis thalia	511	4	36.4	105	21	AAAB27001	Human lambda CL do
439	4	36.4	90	22	AAAB48026	Light chain sequen	512	4	36.4	105	21	AAV51981	HIV-2 ROD isolate
440	4	36.4	91	21	AAAB42378	Human ORFX ORF2142	513	4	36.4	105	21	AAV33248	HIV-2 ROD strain V
441	4	36.4	91	21	AA657450	Arabidopsis thalia	514	4	36.4	106	9	AAAP80814	Sequence of R prot
442	4	36.4	92	12	AAAR15179	hCG/DCG alpha subu	515	4	36.4	106	17	AAAR92516	Microtubule-associ
443	4	36.4	92	21	AAAG02256	Human secreted pro	516	4	36.4	106	18	AAW31727	Alpha light chain
444	4	36.4	93	21	AAV91011	Fowlpox virus 7.3	517	4	36.4	106	19	AAW60941	Streptococcus pneu
445	4	36.4	94	20	AAV28572	Secreted peptide c	518	4	36.4	106	20	AAW70803	Amino acid sequenc
446	4	36.4	95	20	AAW87985	Ubiquitin-like dom	519	4	36.4	106	21	AAAG18736	Zea mays protein f
447	4	36.4	95	20	AAW87986	Ubiquitin-like dom	520	4	36.4	106	21	AAV92193	Human IgG1 lambda
448	4	36.4	95	21	AAAG1103	Zea mays protein f	521	4	36.4	107	19	AAW75895	Peptide inhibitor
449	4	36.4	96	12	AAAR15188	hCG/DCG alpha subu	522	4	36.4	107	19	AAW54013	Anti-Cda antibody

231	4	36.4	9	21	AAB37084	B. brevis HMPp9 P	304	4	36.4	34	20	AAV27769	Human secreted pro
232	4	36.4	10	18	AAW41835	Modified B. burgdo	305	4	36.4	35	19	AAW79008	Rieske motif in KP
233	4	36.4	10	20	AAV46330	Immunogenic peptid	306	4	36.4	35	19	AAW79009	Rieske motif in KP
234	4	36.4	10	21	AAB22031	p67SRF protein fra	307	4	36.4	35	19	AAW79010	Rieske motif in B-
235	4	36.4	10	21	AAB22074	O-glycosylation si	308	4	36.4	35	19	AAW79011	Rieske motif in PP
236	4	36.4	10	21	AAV95865	Bacillus brevis MW	309	4	36.4	35	19	AAW79012	Rieske motif in p5
237	4	36.4	10	21	AAV69786	Listeria innocua P	310	4	36.4	35	19	AAW59701	Amino acid sequenc
238	4	36.4	11	14	AAV54627	Colony stimulating	311	4	36.4	35	21	AAW59466	Arbidolops thalia
239	4	36.4	11	19	AAW41761	Partial sequence #	312	4	36.4	36	20	AAW94500	Bnl brain-specific
240	4	36.4	11	21	AAB26618	SV40 large T antige	313	4	36.4	37	21	AAB32038	Human secreted pro
241	4	36.4	11	21	AAV87645	Lipoprotein derive	314	4	36.4	37	14	AAR39893	Lipopetide TPRV-C
242	4	36.4	12	15	AAR49286	Lipoprotein (a)-pe	315	4	36.4	37	21	AAV94558	Human big endothel
243	4	36.4	12	15	AAR57484	SH2 domain binding	316	4	36.4	37	21	AAB09564	Human Op-1 mutant
244	4	36.4	12	18	AAW25458	SV40 large T antige	317	4	36.4	37	21	AAB09572	Human Op-1 mutant
245	4	36.4	12	21	AAV44719	Anti-contraction p	318	4	36.4	38	21	AAB09575	Human secreted pro
246	4	36.4	13	15	AAR67111	H1A class II bindi	319	4	36.4	39	21	AAW56465	Myobacterium tube
247	4	36.4	13	21	AAV99293	Immunogenic peptid	320	4	36.4	40	19	AAW73052	Yeast Nup2 protein
248	4	36.4	14	22	AAB65584	GVAP fragment p85	321	4	36.4	40	20	AAV08801	Human normal blad
249	4	36.4	14	22	AAR94518	Internal peptide o	322	4	36.4	42	20	AAV60384	polyglutamine trac
250	4	36.4	15	20	AAV43588	Thermophilus them	323	4	36.4	42	22	AAV60384	(1-3)-beta-D-gluc
251	4	36.4	15	20	AAV05912	Human big endothel	324	4	36.4	44	17	AAR93587	Arbidolops thalia
252	4	36.4	16	14	AAR39522	Mouse light chain	325	4	36.4	45	21	AAW38419	Human 5' EST seque
253	4	36.4	16	15	AAR52128	Mouse light chain	326	4	36.4	48	20	AAV12839	Human secreted pro
254	4	36.4	16	15	AAR52129	Mouse light chain	327	4	36.4	48	21	AAB45048	Human secreted pro
255	4	36.4	16	15	AAR52130	Mouse light chain	328	4	36.4	48	21	AAW54856	Arbidolops thalia
256	4	36.4	16	15	AAR52131	Mouse light chain	329	4	36.4	48	21	AAW54856	Arbidolops thalia
257	4	36.4	16	15	AAR52132	Mouse light chain	330	4	36.4	48	21	AAW57356	Arbidolops thalia
258	4	36.4	16	15	AAR52133	Mouse light chain	331	4	36.4	48	21	AAW60514	Arbidolops thalia
259	4	36.4	16	15	AAR52135	Mouse light chain	332	4	36.4	49	19	AAW74806	Human secreted pro
260	4	36.4	16	15	AAR52135	Mouse light chain	333	4	36.4	50	20	AAV36060	Extended human sec
261	4	36.4	16	15	AAR52135	Mouse light chain	334	4	36.4	50	20	AAV01152	Secreted protein e
262	4	36.4	16	15	AAR52124	Interferon-gamma r	335	4	36.4	51	21	AAW07981	Arbidolops thalia
263	4	36.4	16	18	AAR98414	Peptide resembling	336	4	36.4	51	21	AAW07981	Fragment of human
264	4	36.4	16	18	AAW38968	HTLV-II gag 4-20 d	337	4	36.4	52	20	AAV02982	Human cancer assoc
265	4	36.4	17	11	AAR06876	Leucine aminopepti	338	4	36.4	52	21	AAB43415	Bacteriophage Dp-1
266	4	36.4	17	13	AAR28142	Epsilon-sarcoglyca	339	4	36.4	52	21	AAB16818	Arbidolops thalia
267	4	36.4	17	21	AAV90743	Kex2 protease inh	340	4	36.4	52	21	AAW57355	Arbidolops thalia
268	4	36.4	18	19	AAW41423	GVAP tyryptic pept	341	4	36.4	52	21	AAW60513	Arbidolops thalia
269	4	36.4	19	17	AAR94525	GVAP fragment p85	342	4	36.4	52	21	AAW02040	Human secreted pro
270	4	36.4	19	17	AAR94510	Human ENOS peptide	343	4	36.4	52	21	AAW02041	Human secreted pro
271	4	36.4	19	19	AAW81232	Human INOS peptide	344	4	36.4	52	21	AAW03409	Human secreted pro
272	4	36.4	19	19	AAW81232	Kex2 protease inh	345	4	36.4	53	13	AAR25038	Sequence encoded b
273	4	36.4	19	19	AAW41428	Internal peptide o	346	4	36.4	53	21	AAW03429	Human secreted pro
274	4	36.4	19	20	AAV43587	Internal peptide o	347	4	36.4	53	22	AAB37486	AGP-1 protein. S
275	4	36.4	19	20	AAV43595	Yada homologous pe	348	4	36.4	54	15	AAR58335	Hypotensive polype
276	4	36.4	19	21	AAB51600	Kentucky Blue Gras	349	4	36.4	54	20	AAV13143	Human secreted pro
277	4	36.4	20	17	AAR89385	Peptide derived fr	350	4	36.4	54	21	AAW01724	Human secreted pro
278	4	36.4	20	20	AAV01694	Hepatitis GB virus	351	4	36.4	55	21	AAW40992	Zea mays protein f
279	4	36.4	20	21	AAB09169	House dust mite al	352	4	36.4	55	17	AAR89417	WC1 N-terminal pe
280	4	36.4	20	22	AAV52511	PL peptide #31. S	353	4	36.4	56	19	AAW38491	S. pneumoniae pne
281	4	36.4	20	22	AAB55820	Syndecan-2 PL pept	354	4	36.4	56	20	AAV50462	Human normal blad
282	4	36.4	20	22	AAB57682	Syndecan-2 PL pept	355	4	36.4	58	19	AAW71565	Hepatocyte nuclear
283	4	36.4	20	22	AAB58092	Proteobacterial ex	356	4	36.4	58	19	AAV20977	Fragment of human
284	4	36.4	21	21	AAB51528	Listeria p60 pepti	357	4	36.4	59	20	AAV20977	Human glial fibrin
285	4	36.4	23	19	AAR45167	Human neuroinflame	358	4	36.4	60	21	AAB38161	Human secreted pro
286	4	36.4	23	19	AAV20505	Rhinovirus genome-	359	4	36.4	60	21	AAW02638	Human secreted pro
287	4	36.4	23	22	AAB73023	Mouse heavy chain	360	4	36.4	60	21	AAW02638	Human secreted pro
288	4	36.4	26	15	AAR52411	Mouse heavy chain	361	4	36.4	62	21	AAR89416	Human secreted pro
289	4	36.4	26	15	AAR52407	Mouse heavy chain	362	4	36.4	62	17	AAR89416	Human secreted pro
290	4	36.4	26	15	AAR52381	Mouse heavy chain	363	4	36.4	62	17	AAW26875	WC1 N-terminal. pe
291	4	36.4	26	15	AAR52360	Mouse heavy chain	364	4	36.4	63	17	AAW26875	Zea mays protein f
292	4	36.4	26	15	AAR51725	Neurotrophic facto	365	4	36.4	63	17	AAW26875	Arbidolops thalia
293	4	36.4	27	18	AAW01786	Residual protease	366	4	36.4	64	14	AAR37963	P. aeruginosa opri
294	4	36.4	27	20	AAW78130	Human secreted pro	367	4	36.4	64	20	AAV73836	KGPR Ig-1-like dom
295	4	36.4	28	21	AAV93257	Amino acid sequenc	368	4	36.4	64	20	AAV73836	Human endometrium
296	4	36.4	29	20	AAV41512	Fragment of human	369	4	36.4	65	20	AAV27640	Human secreted pro
297	4	36.4	30	15	AAR47193	N-terminal fragmen	370	4	36.4	66	21	AAW16868	Arbidolops thalia
298	4	36.4	30	22	AAB75250	Human secreted pro	371	4	36.4	66	21	AAW40982	Zea mays protein f
299	4	36.4	31	14	AAR34860	Candida acidic pro	372	4	36.4	69	14	AAR44553	v1 protein of fel
300	4	36.4	31	14	AAR34861	Candida acidic pro	373	4	36.4	69	20	AAV19503	Amino acid sequenc
301	4	36.4	32	21	AAB07999	Peptide derived fr	374	4	36.4	69	21	AAW21730	Arbidolops thalia
302	4	36.4	33	21	AAB39405	Human secreted pro	375	4	36.4	70	19	AAW22897	Seq ID NO. 93 from
303	4	36.4	34	14	AAR39873	C peptide RV-C6, r	376	4	36.4	70	20	AAW67845	Human secreted pro

85	5	45.5	279	17	AAW0803	Thermilase L221Q.	158	5	45.5	635	21	AAV70781	EGFP-VASP fusion p
86	5	45.5	279	17	AAW0769	Thermilase T111N.	159	5	45.5	638	11	AAAR03924	E. coli HSP (dnak)
87	5	45.5	279	17	AAW00770	Thermilase T217G +	160	5	45.5	668	21	AAAB57099	Human prostrate can
88	5	45.5	279	17	AAW00771	Thermilase G139N.	161	5	45.5	668	21	AAAB28595	Maize Scarcrow pr
89	5	45.5	279	17	AAW62226	Subtilase TWYHER F	162	5	45.5	714	19	AAW98361	H. pylori GHPO 137
90	5	45.5	279	20	AAI24912	Subtilase subcll1	163	5	45.5	723	21	AAV75477	Neisseria meningit
91	5	45.5	279	20	AAW21650	Subtilase TWYHER.	164	5	45.5	730	17	AAV43385	S. pneumoniae PspC
92	5	45.5	279	20	AAW92816	T. thalophilus TH	165	5	45.5	746	17	AAAR3270	GST-INL fusion int
93	5	45.5	280	21	AAAB3251	Streptomyces coll.	166	5	45.5	756	21	AAAB2605	Human ORF ORF2369
94	5	45.5	292	13	AAAR29312	Cyclin D3 protein.	167	5	45.5	757	21	AAV44364	Human cell cycle r
95	5	45.5	292	14	AAAR4804	Human cyclin D3.	168	5	45.5	758	21	AAAB3319	Human colon cancer
96	5	45.5	296	17	AAW04577	Sensory and motor	169	5	45.5	766	20	AAV13457	Human acid sequenc
97	5	45.5	296	17	AAAR97224	Sensory and motor	170	5	45.5	770	20	AAV13456	Amino acid sequenc
98	5	45.5	296	19	AAW41263	Sensory and motor	171	5	45.5	783	20	AAV60344	Human normal blad
99	5	45.5	296	19	AAW46465	Sensory and motor	172	5	45.5	818	19	AAW72035	HSV-2 strain SB5 C
100	5	45.5	296	20	AAV06640	Sensory and motor	173	5	45.5	859	20	AAV00991	Human ATP-depend
101	5	45.5	296	21	AAV71177	Human Sensory and	174	5	45.5	859	22	AAAB37463	Human ATP-depend
102	5	45.5	296	21	AAV57385	Sensory and motor	175	5	45.5	888	18	AAW16494	DNA polymerase whl
103	5	45.5	302	19	AAW72214	HSV-2 strain SB5 C	176	5	45.5	950	20	AAV33298	Human membrane spa
104	5	45.5	320	18	AAW19990	Type I, p80 IL-1-r	177	5	45.5	1004	21	AAAB1841	Haemophilus influe
105	5	45.5	320	15	AAAB37795	Human Interleukin-	178	5	45.5	1010	21	AAAB1840	Haemophilus influe
106	5	45.5	322	15	AAAB52025	Bacillus sp. NKS-2	179	5	45.5	1050	21	AAAB99774	Molecular pathogen
107	5	45.5	323	20	AAV37286	Amino acid sequenc	180	5	45.5	1051	20	AAAB9741	A serine/threonine
108	5	45.5	323	22	AAAB9174	Human G-protein co	181	5	45.5	1061	21	AAAB80568	Wild type Escheric
109	5	45.5	327	21	AAAB33321	Eucalyptus grandis	182	5	45.5	1070	13	AAAR21521	Alpha galactosidas
110	5	45.5	339	16	AAAR77317	Protein activated	183	5	45.5	1070	22	AAAB48099	Amino acid sequenc
111	5	45.5	342	22	AAAB48740	Mouse liver growth	184	5	45.5	1116	21	AAAB2449	Human ORF ORF2213
112	5	45.5	343	14	AAAR41670	Porcine stalytiran	185	5	45.5	1237	19	AAAM44729	Chicken protein ty
113	5	45.5	343	16	AAAB65240	L. lactis branched	186	5	45.5	1237	20	AAAB9347	Chicken transmembr
114	5	45.5	344	15	AAAR54222	Human prostrate can	187	5	45.5	1257	17	AAAB87628	Alpha-ketoglutaric
115	5	45.5	344	21	AAAB56576	Human breast cance	188	5	45.5	1257	19	AAW41781	B. lactofermentum
116	5	45.5	347	22	AAAB63245	Human death associ	189	5	45.5	1339	21	AAAB42887	Human ORF ORF2651
117	5	45.5	373	21	AAAG04115	Arabidopsis thalia	190	5	45.5	1522	20	AAV03183	Topoisomerase II b
118	5	45.5	373	21	AAAG1856	Arabidopsis thalia	191	5	45.5	1634	14	AAAR42452	Enzyme involved in
119	5	45.5	373	21	AAAG54002	Arabidopsis thalia	192	5	45.5	1642	20	AAAB9400	S. putrefaciens PK
120	5	45.5	398	16	AAAR74206	Human death associ	193	5	45.5	1876	18	AAW38757	Phosphatidylyl inosl
121	5	45.5	398	19	AAAR42368	Death associated p	194	5	45.5	1896	19	AAW20995	HSV-2 strain SB5 C
122	5	45.5	398	21	AAAB42325	Human ORF ORF2089	195	5	45.5	1920	22	AAAB65656	Novel protein kina
123	5	45.5	415	21	AAV77273	Streptomyces viola	196	5	45.5	2023	21	AAV54330	Amino acid sequenc
124	5	45.5	415	21	AAV78825	Amino acid sequenc	197	5	45.5	2074	21	AAV54319	Amino acid sequenc
125	5	45.5	429	21	AAV74547	Neisseria meningit	198	5	45.5	2135	22	AAU00019	Human Plexin prote
126	5	45.5	429	21	AAV74548	Neisseria meningit	199	5	45.5	2756	17	AAAR99462	Biosynthetic enzym
127	5	45.5	435	21	AAAG36200	Arabidopsis thalia	200	5	45.5	2756	20	AAW37050	S. putrefaciens BP
128	5	45.5	438	15	AAAR47259	Pre-pro-VGR1. Mus	201	5	45.5	2756	21	AAAB10467	Shewanella putrefe
129	5	45.5	449	22	AAAB48100	Neisseria acid sequenc	202	5	45.5	3084	10	AAAP94758	Sequence of mouse
130	5	45.5	451	20	AAV38573	Neisseria meningit	203	5	45.5	3119	19	AAAR72204	HSV-2 strain SB5 C
131	5	45.5	451	20	AAV38574	DAX-1 protein. Ho	204	5	45.5	3266	21	AAAB42491	Human ORF ORF2255
132	5	45.5	470	18	AAAI2678	CRD domain of huma	205	5	36.4	7	16	AAV49176	C-terminal domain
133	5	45.5	478	22	AAAB48101	Mouse CRD amino ac	206	5	36.4	7	16	AAAR81831	Synthetic peptide
134	5	45.5	486	22	AAAB48104	Recombinant PDI (A	207	5	36.4	7	17	AAAR81832	Insulin production
135	5	45.5	491	13	AAAR25296	PDI. Homo sapiens	208	5	36.4	7	20	AAW07158	Plasminogen activa
136	5	45.5	508	9	AAAB0664	Human PDI. Homo s	209	5	36.4	8	16	AAV50559	Amino acid sequenc
137	5	45.5	508	13	AAAR25287	Bovine protein dis	210	5	36.4	8	21	AAAB21394	Somatic immunoglob
138	5	45.5	508	13	AAAR25287	Human PDI. Homo s	211	5	36.4	8	16	AAAB21394	Listeria p60 pepti
139	5	45.5	510	9	AAAB0615	Amino acid sequenc	212	5	36.4	8	21	AAAB29834	P. faicicparum lact
140	5	45.5	513	22	AAAB31008	B. napus FSH poly	213	5	36.4	9	11	AAAR84506	Peptide (38) inh
141	5	45.5	513	22	AAAB48181	Arabidopsis thalia	214	5	36.4	9	14	AAAR59945	Peptide (68) inh
142	5	45.5	515	13	AAAR25298	Human HSA-PDI fusi	215	5	36.4	9	15	AAAR57568	Peptide (146) inh
143	5	45.5	515	15	AAAR51697	Arabidopsis thalia	216	5	36.4	9	15	AAAR57568	Peptide (140) inh
144	5	45.5	519	21	AAAG36199	Arabidopsis thalia	217	5	36.4	9	15	AAAR57568	Peptide (104) inh
145	5	45.5	520	18	AAAR26640	Arabidopsis thalia	218	5	36.4	9	15	AAAR57568	Peptide (110) inh
146	5	45.5	520	19	AAW40099	Arabidopsis ferula	219	5	36.4	9	15	AAAR57568	Peptide (237) inh
147	5	45.5	520	21	AAAB15188	Amino acid sequenc	220	5	36.4	9	15	AAAR57568	Peptide (211) inh
148	5	45.5	520	22	AAAB31007	B. napus FSH poly	221	5	36.4	9	15	AAAR57568	Peptide (181) inh
149	5	45.5	520	22	AAAB48179	B. napus FSH poly	222	5	36.4	9	15	AAAR57568	Peptide (146) inh
150	5	45.5	520	22	AAAB48180	Arabidopsis thalia	223	5	36.4	9	15	AAAR57568	Peptide (176) inh
151	5	45.5	521	21	AAAG31483	Arabidopsis thalia	224	5	36.4	9	15	AAAR57568	Peptide (140) inh
152	5	45.5	522	21	AAAG31483	Tomato phosphate t	225	5	36.4	9	15	AAAR57568	Peptide (110) inh
153	5	45.5	528	19	AAW46750	Arabidopsis thalia	226	5	36.4	9	15	AAAR57568	Peptide (237) inh
154	5	45.5	537	21	AAAG31482	Arabidopsis thalia	227	5	36.4	9	15	AAAR57568	Synthetic peptide
155	5	45.5	540	21	AAAG31481	A serine/threonine	228	5	36.4	9	15	AAAR57568	Peptide derived fr
156	5	45.5	582	20	AAW67659	Green fluorescent	229	5	36.4	9	17	AAV01696	
157	5	45.5	635	19	AAW65034		230	5	36.4	9	20	AAV01696	

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:35:31 ; Search time 56.84 Seconds  
(without alignments)  
11,732 Million cell updates/sec

Title: US-09-372-036-29  
Perfect score: 11  
Sequence: 1 QOOTAPKAPRE 11

Scoring table:  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

\_A\_Geneseq\_0601.\*  
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	11	100.0	11	AA45171	Listeria p60 pepti
2	11	100.0	21	AA473894	Listeria monocytog
3	11	100.0	478	AA45178	Listeria p60 prote
4	11	100.0	484	AA473913	Listeria monocytog
5	9	81.8	9	AA454618	Listeria p60 pepti
6	9	81.8	23	AA45162	Listeria p60 pepti
7	7	63.6	7	AA473906	Listeria monocytog
8	6	54.5	150	AA426181	Rice acetonydroxya
9	6	54.5	890	AA460252	NF-ATP. Mus sp.
10	5	45.5	11	AA454621	Listeria p60 pepti
11	5	45.5	13	AA454625	Listeria innocua p

12	5	45.5	25	14	AA45165
13	5	45.5	54	21	AA494958
14	5	45.5	95	20	AA460235
15	5	45.5	96	20	AA429500
16	5	45.5	96	21	AA444435
17	5	45.5	107	19	AA457218
18	5	45.5	107	19	AA457219
19	5	45.5	107	19	AA457220
20	5	45.5	107	19	AA457221
21	5	45.5	107	21	AA457221
22	5	45.5	107	21	AA457221
23	5	45.5	107	21	AA457221
24	5	45.5	107	21	AA457221
25	5	45.5	107	21	AA457221
26	5	45.5	107	21	AA457221
27	5	45.5	107	21	AA457221
28	5	45.5	107	21	AA457221
29	5	45.5	107	21	AA457221
30	5	45.5	107	21	AA457221
31	5	45.5	107	21	AA457221
32	5	45.5	107	21	AA457221
33	5	45.5	107	21	AA457221
34	5	45.5	107	21	AA457221
35	5	45.5	107	21	AA457221
36	5	45.5	107	21	AA457221
37	5	45.5	107	21	AA457221
38	5	45.5	107	21	AA457221
39	5	45.5	107	21	AA457221
40	5	45.5	107	21	AA457221
41	5	45.5	107	21	AA457221
42	5	45.5	107	21	AA457221
43	5	45.5	107	21	AA457221
44	5	45.5	107	21	AA457221
45	5	45.5	107	21	AA457221
46	5	45.5	107	21	AA457221
47	5	45.5	107	21	AA457221
48	5	45.5	107	21	AA457221
49	5	45.5	107	21	AA457221
50	5	45.5	107	21	AA457221
51	5	45.5	107	21	AA457221
52	5	45.5	107	21	AA457221
53	5	45.5	107	21	AA457221
54	5	45.5	107	21	AA457221
55	5	45.5	107	21	AA457221
56	5	45.5	107	21	AA457221
57	5	45.5	107	21	AA457221
58	5	45.5	107	21	AA457221
59	5	45.5	107	21	AA457221
60	5	45.5	107	21	AA457221
61	5	45.5	107	21	AA457221
62	5	45.5	107	21	AA457221
63	5	45.5	107	21	AA457221
64	5	45.5	107	21	AA457221
65	5	45.5	107	21	AA457221
66	5	45.5	107	21	AA457221
67	5	45.5	107	21	AA457221
68	5	45.5	107	21	AA457221
69	5	45.5	107	21	AA457221
70	5	45.5	107	21	AA457221
71	5	45.5	107	21	AA457221
72	5	45.5	107	21	AA457221
73	5	45.5	107	21	AA457221
74	5	45.5	107	21	AA457221
75	5	45.5	107	21	AA457221
76	5	45.5	107	21	AA457221
77	5	45.5	107	21	AA457221
78	5	45.5	107	21	AA457221
79	5	45.5	107	21	AA457221
80	5	45.5	107	21	AA457221
81	5	45.5	107	21	AA457221
82	5	45.5	107	21	AA457221
83	5	45.5	107	21	AA457221
84	5	45.5	107	21	AA457221

Listeria p60 pepti  
Human secreted pro  
Human endometrium  
Human lung tumour  
Human lung tumour  
Desulfovibrio vulg  
Cytochrome c3 prot  
Cytochrome c3 prot  
Cytochrome c3 prot  
D. vulgaris cytoch  
Mutant cytochrome  
Mutant cytochrome  
Mutant cytochrome  
Wild type cytochro  
Mutant cytochrome  
M. tuberculosis an  
M. tuberculosis re  
Pneumococcal sur  
Chlamydia pneumoni  
Human secreted pro  
Human secreted pro  
Polypeptide isolat  
Hepatitis GB virus  
Pseudomonas syring  
Hepatitis GB virus  
Human breast cance  
Corynebacterium gl  
E. coli RNase E.  
Escherichia coli R  
Group B Streptococ  
Human ORF ORP2661  
Streptococcus pneu  
S. pneumoniae prot  
Sequence of sublin  
pG4-5-CDK-BP clon  
Amino acid sequenc  
Thermilase (p229X)  
Thermilase variant  
Thermilase P205Q,  
Thermilase L104D,  
Thermilase Q66N,  
Thermilase P205G,  
Thermilase G206N,  
Thermilase N68D, V  
Thermilase N66S, S  
Thermilase S109E,  
Thermilase S109E,  
Thermilase S207E,  
Thermilase L134D,  
Thermilase A168N,  
Thermilase G135Q,  
Thermilase Y210G,  
Thermilase T195P,  
Thermilase N166D,  
Thermilase W208F,  
Thermilase G67E, S  
Thermilase S191E,  
Thermilase Y171T,  
Thermilase S192D,  
Thermilase I209A,  
Thermilase Y216N,  
Thermilase L216N,  
Thermilase Y218A,  
Thermilase W112C,  
Thermilase V103A,  
Thermilase A164H,  
Thermilase L221V,  
Thermilase S191D,  
Thermilase G135Q,



APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1502  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-412-227-29

Query Match 100.0%; Score 57; DB 8; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.006;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
|||||  
Db 1 000TAPKAPTE 11

RESULT 2  
US-08-412-227A-29  
Sequence 29, Application US/08412227A  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/412,227A  
FILING DATE: 27-MAR-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4

FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-412-227A-29

Query Match 100.0%; Score 57; DB 8; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.006;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
|||||  
Db 1 000TAPKAPTE 11

RESULT 3  
US-08-412-227C-29  
Sequence 29, Application US/08412227C  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/412,227C  
FILING DATE: 27-MAR-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4



FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-412-227C-29

Query Match 100.0%; Score 57; DB 8; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.006;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
|||||  
Db 1 000TAPKAPTE 11

## RESULT 4

US-09-372-036-29

Sequence 29, Application US/09372036

GENERAL INFORMATION:

APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/372,036

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/456,670

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/075,248

FILING DATE: 11-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4239567.4

FILING DATE: 25-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4219111.4

FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-09-372-036-29

Query Match 100.0%; Score 57; DB 17; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.006;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
|||||  
Db 1 000TAPKAPTE 11

## RESULT 5

US-08-412-227-42

Sequence 42, Application US/08412227

GENERAL INFORMATION:

APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/412,227

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/075,248

FILING DATE: 11-JUN-1993

APPLICATION NUMBER: DE 4239567.4

FILING DATE: 25-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4219111.4

FILING DATE: 11-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: HAMLET-KING, DIANA

REGISTRATION NUMBER: 33,302

REFERENCE/DOCKET NUMBER: MERCK 1502  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-412-227-42

Query Match 100.0%; Score 57; DB 8; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0066;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOOTAPKAPTE 11  
Db 2 QOOTAPKAPTE 12

RESULT 6  
US-08-412-227A-42  
Sequence 42, Application US/08412227A  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINXWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STERN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/412,227A  
FILING DATE: 27-MAR-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-412-227C-42

SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-412-227A-42

Query Match 100.0%; Score 57; DB 8; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0066;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOOTAPKAPTE 11  
Db 2 QOOTAPKAPTE 12

RESULT 7  
US-08-412-227C-42  
Sequence 42, Application US/08412227C  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINXWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STERN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/412,227C  
FILING DATE: 27-MAR-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-412-227C-42

Query Match 100.0%; Score 57; DB 8; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0066;  
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
111111111111  
DB 2 000TAPKAPTE 12

RESULT 8  
US-09-372-036-42  
Sequence 42, Application US/09372036

GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, STEFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, COTTFRIED  
APPLICANT: BOBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANTIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/372.036  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/456,670  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-372-036-42

Query Match 100.0%; Score 57; DB 17; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.0066;  
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
111111111111  
DB 2 000TAPKAPTE 12

RESULT 9  
US-08-479-520-33  
Sequence 33, Application US/08479520

GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THERETO, AND USES THEREOF  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/479,520  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/105/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-479-520-33

Query Match 100.0%; Score 57; DB 8; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 11: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
111111111111  
DB 7 000TAPKAPTE 17

RESULT 10  
US-08-486-050-33  
Sequence 33, Application US/08486050

GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THERETO, AND USES THEREOF  
NUMBER OF SEQUENCES: 75

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Foley & Lardner  
;; STREET: 3000 K Street, N.W., Suite 500  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20007-5109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/486,050  
;; FILING DATE: 07-JUN-1995  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/127,499  
;; FILING DATE: 28-SEP-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BENT, Stephen A.  
;; REGISTRATION NUMBER: 29,768  
;; REFERENCE/DOCKET NUMBER: 51916/103/INBI  
;; TELEPHONE: (202)672-5300  
;; TELEFAX: (202)672-5399  
;; TELEX: 904136  
;; INFORMATION FOR SEQ ID NO: 33:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 21 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: unknown  
;; US-08-486-050-33

Query Match 100.0%; Score 57; DB 8; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
DB 7 000TAPKAPTE 17

RESULT 11  
US-08-486-050A-33  
;; Sequence 33, Application US/08486050A  
;; GENERAL INFORMATION:  
;; APPLICANT: VAN ALSTYNE, Diane  
;; APPLICANT: SHARMA, Lawrence Rajendra  
;; TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
;; TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
;; TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
;; NUMBER OF SEQUENCES: 75  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Foley & Lardner  
;; STREET: 3000 K Street, N.W., Suite 500  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20007-5109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/486,050A  
;; FILING DATE: 07-JUN-1995  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/127,499

;; FILING DATE: 28-SEP-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BENT, Stephen A.  
;; REGISTRATION NUMBER: 29,768  
;; REFERENCE/DOCKET NUMBER: 51916/103/INBI  
;; TELEPHONE: (202)672-5300  
;; TELEFAX: (202)672-5399  
;; TELEX: 904136  
;; INFORMATION FOR SEQ ID NO: 33:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 21 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: unknown  
;; US-08-486-050A-33

Query Match 100.0%; Score 57; DB 8; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 000TAPKAPTE 11  
DB 7 000TAPKAPTE 17

RESULT 12  
US-08-988-444-33  
;; Sequence 33, Application US/08988444  
;; GENERAL INFORMATION:  
;; APPLICANT: VAN ALSTYNE, Diane  
;; APPLICANT: SHARMA, Lawrence Rajendra  
;; TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
;; TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
;; TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
;; NUMBER OF SEQUENCES: 75  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Foley & Lardner  
;; STREET: 3000 K Street, N.W., Suite 500  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: USA  
;; ZIP: 20007-5109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/988,444  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/486,050  
;; FILING DATE: 07-JUN-1995  
;; APPLICATION NUMBER: US 08/127,499  
;; FILING DATE: 28-SEP-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: BENT, Stephen A.  
;; REGISTRATION NUMBER: 29,768  
;; REFERENCE/DOCKET NUMBER: 51916/103/INBI  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202)672-5300  
;; TELEFAX: (202)672-5399  
;; TELEX: 904136  
;; INFORMATION FOR SEQ ID NO: 33:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 21 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: unknown  
;; US-08-988-444-33

Query Match 100.0%; Score 57; DB 13; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OOOTAPKAPTE 11  
 |||||  
 Db 7 OOOTAPKAPTE 17

## RESULT 13

US-09-217-849-33  
 : Sequence 33, Application US/09217849

## : GENERAL INFORMATION:

APPLICANT: VAN ALSTYNE, Diane  
 TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
 TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
 TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THERETO, AND USES THEREOF  
 NUMBER OF SEQUENCES: 75

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner  
 STREET: 3000 K Street, N.W., Suite 500  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA

ZIP: 20007-5109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/217,849

FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/486,050

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/127,499

FILING DATE: 28-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 51916/103/INBI

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

US-09-217-849-33

Query Match 100.0%; Score 57; DB 16; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 0.012;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OOOTAPKAPTE 11  
 |||||  
 Db 7 OOOTAPKAPTE 17

## RESULT 14

US-08-412-227-39

: Sequence 39, Application US/08412227

: GENERAL INFORMATION:

: APPLICANT: SCHUBERT, PETER

APPLICANT: NEUMANN, SIEGFRIED  
 APPLICANT: PAMELZIK, MARTINA  
 APPLICANT: LINKEWILLER, WINFRIED

APPLICANT: BURGER, CHRISTA

APPLICANT: HOFMANN, GOTTFRIED

APPLICANT: HUBERT, ANDREAS

APPLICANT: GOEBEL, WERNER

APPLICANT: KOHLER, STEFAN

TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING

TITLE OF INVENTION: LISTERIAS

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.

STREET: 2200 CLARENDON BLVD., SUITE 1400

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: US

ZIP: 22201

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/412,227

FILING DATE:

## CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/075,248

FILING DATE: 11-JUN-1993

APPLICATION NUMBER: DE 4239567.4

FILING DATE: 25-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4219111.4

FILING DATE: 11-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: HAMLET-KING, DIANA

REGISTRATION NUMBER: 33,302

REFERENCE/DOCKET NUMBER: MERCK 1502

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-243-6333

TELEFAX: 703-243-6410

TELEX: 64191

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 232 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORGANISM: Listeria monocytogenes

STRAIN: EGD

US-08-412-227-39

Query Match 100.0%; Score 57; DB 8; Length 232;  
 Best Local Similarity 100.0%; Pred. No. 0.19;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 OOOTAPKAPTE 11  
 |||||  
 Db 222 OOOTAPKAPTE 232

## RESULT 15

US-08-412-227A-39

: Sequence 39, Application US/08412227A

: GENERAL INFORMATION:

: APPLICANT: SCHUBERT, PETER

: APPLICANT: NEUMANN, SIEGFRIED

: APPLICANT: PAMELZIK, MARTINA

: APPLICANT: LINKEWILLER, WINFRIED

: APPLICANT: BURGER, CHRISTA

APPLICANT: HOEWANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/412,227A  
FILING DATE: 27-MAR-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 232 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-412-227A-39

Query Match 100.0%; Score 57; DB 8; Length 232;  
Best Local Similarity 100.0%; Pred. No. 0.19;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QOOTAPKAPTE 11  
|||||  
DB 222 QOOTAPKAPTE 232

Search completed: August 15, 2001, 13:15:54  
Job time: 391 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 13:10:30 ; Search time 21.63 seconds  
(without alignments)  
7.948 Million cell updates/sec

Title: US-09-372-036-29

Perfect score: 57

Sequence: 1 OQOTAPKAPTE 11

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 109802 seqs, 15628634 residues

Total number of hits satisfying chosen parameters: 109802

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 0%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCr\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	100.0	21	5	US-09-489-850-33
2	57	100.0	484	5	US-09-489-850-26
3	36	63.2	100	5	US-09-803-110-9462
4	36	63.2	410	5	US-09-760-443-1455
5	35	61.4	7	5	US-09-489-850-34
6	35	61.4	198	5	US-09-758-459-219
7	35	61.4	310	1	PCr-US01-16450-3703
8	35	61.4	353	1	PCr-US01-18569-3045
9	35	61.4	446	5	US-09-909-446-2
10	35	61.4	2626	5	US-09-477-962-106
11	34	59.6	120	5	US-09-758-468-419
12	34	59.6	239	1	PCr-US01-14827-14067
13	34	59.6	240	5	US-09-758-470-399
14	34	59.6	441	5	US-09-760-462-27
15	33	57.9	143	1	PCr-US01-08656-7761
16	33	57.9	147	5	US-09-758-471-3699
17	33	57.9	211	5	US-09-764-870-414
18	33	57.9	303	1	PCr-US01-04926A-326
19	33	57.9	358	5	US-09-803-110-12650
20	33	57.9	372	1	PCr-US01-04926A-325
21	33	57.9	414	5	US-09-758-471-3744
22	33	57.9	719	1	PCr-US01-14827-14710
23	33	57.9	964	5	US-09-760-469-1491
24	32	56.1	95	5	US-09-760-444-285
25	32	56.1	125	5	US-09-764-847-883
26	32	56.1	180	1	PCr-US01-14827-9696
27	32	56.1	298	5	US-09-765-272-24

28	32	56.1	328	5	US-09-891-126-5	Sequence 5, App1
29	32	56.1	483	5	US-09-760-466-1241	Sequence 1241, Ap
30	32	56.1	1074	1	PCr-US01-08656-10235	Sequence 10235, A
31	32	56.1	1162	5	US-09-894-273-2	Sequence 2, App1
32	32	56.1	1418	5	PCr-US01-08656-10236	Sequence 10236, A
33	31.5	55.3	780	5	US-09-803-110-13189	Sequence 13189, A
34	31	54.4	48	5	US-09-757-028-2364	Sequence 2364, Ap
35	31	54.4	79	1	PCr-US01-08656-7604	Sequence 7604, Ap
36	31	54.4	96	5	US-09-760-498-704	Sequence 704, App
37	31	54.4	115	5	US-09-738-469-681	Sequence 681, App
38	31	54.4	117	5	US-09-758-472-8918	Sequence 8918, Ap
39	31	54.4	135	5	US-09-758-472-7896	Sequence 7896, Ap
40	31	54.4	138	5	US-09-758-459-246	Sequence 246, App
41	31	54.4	141	1	PCr-US01-16450-1928	Sequence 1928, Ap
42	31	54.4	144	5	US-09-764-877-1818	Sequence 1818, Ap
43	31	54.4	166	1	PCr-US01-14827-13722	Sequence 13722, A
44	31	54.4	172	1	PCr-US01-08656-8740	Sequence 8740, Ap
45	31	54.4	180	5	US-09-758-466-670	Sequence 670, App

#### ALIGNMENTS

RESULT 1  
US-09-489-850-33  
Sequence 33, Application US/09489850  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/489, 850  
FILING DATE: 24-Jan-2000  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/988,444  
FILING DATE: <unknown>  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENNETT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/103/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: <unknown>  
TOPOLOGY: unknown  
SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-09-489-850-33  
Query Match 100.0%; Score 57; DB 5; Length 21;





TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/489,850  
FILING DATE: 24-Jan-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/988,444  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/103/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: unknown  
SEQUENCE DESCRIPTION: SEQ ID NO: 34:  
US-09-489-850-34  
Query Match 61.4%; Score 35; DB 5; Length 7;  
Best Local Similarity 100.0%; Pred. No. 9.4e+04;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 OCTAPKA 8  
DB 1 OCTAPKA 7  
RESULT 6  
US-09-758-459-219  
; Sequence 219, Application US/09758459  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PM029  
; CURRENT APPLICATION NUMBER: US/09/758,459  
; CURRENT FILING DATE: 2001-01-11  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; NUMBER OF SEQ ID NOS: 406  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 219  
; LENGTH: 198  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-459-219

Query Match 61.4%; Score 35; DB 5; Length 198;  
Best Local Similarity 75.0%; Pred. No. 40;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 4 TAPKAPTE 11  
DB 161 TAPESPTE 168  
RESULT 7  
PCT-US01-16450-2703  
; Sequence 2703, Application PC/TUS0116450  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA131PCT  
; CURRENT APPLICATION NUMBER: PCT/US01/16450  
; CURRENT FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2820  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2703  
; LENGTH: 310  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US01-16450-2703

Query Match 61.4%; Score 35; DB 1; Length 310;  
Best Local Similarity 75.0%; Pred. No. 65;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 4 TAPKAPTE 11  
DB 191 TAPESPTE 198

RESULT 8  
PCT-US01-18569-3045  
; Sequence 3045, Application PC/TUS0118569  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA133PCT  
; CURRENT APPLICATION NUMBER: PCT/US01/18569  
; CURRENT FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 4360  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3045  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US01-18569-3045

Query Match 61.4%; Score 35; DB 1; Length 353;  
Best Local Similarity 66.7%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 3 OTAPKAPTE 11  
DB 53 ETAPRAPRE 61

RESULT 9  
US-09-909-446-2  
; Sequence 2, Application US/09909446  
; GENERAL INFORMATION:  
; APPLICANT: ENMARK, EVA

```

; GUSTAFSSON, JAN
; TITLE OF INVENTION: OR-1 ON ORPHAN RECEPTOR BELONGING
; TO THE NUCLEAR RECEPTOR FAMILY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Banner & Witcoff
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/909,446
; FILING DATE: 19-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/776,844
; FILING DATE: <Unknown>
; APPLICATION NUMBER: UK 9413536.2
; FILING DATE: 16-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A
; REGISTRATION NUMBER: 32141
; REFERENCE/DOCKET NUMBER: 00487.04029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 446 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-909-446-2

Query Match          61.4%: Score 35; DB 5; Length 446;
Best Local Similarity 63.6%: Pred. No. 98;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 QQTAPKAPTE 11
    111 1 111
Db 164 QQQQQPPPT 174

RESULT 10
US-09-477-962-106
; Sequence 106, Application US/09477962
; GENERAL INFORMATION:
; APPLICANT: SHEN, BEN
; APPLICANT: DU, LIANGCHENG
; APPLICANT: SANCHEZ, CESAR
; APPLICANT: CHEN, MEI
; APPLICANT: EDWARDS, DANIEL J.
; TITLE OF INVENTION: BLEOMYCIN GENE CLUSTER COMPONENTS AND THEIR USES
; FILE REFERENCE: 407T-895820US
; CURRENT APPLICATION NUMBER: US/09/477,962
; CURRENT FILING DATE: 2000-01-05
; PRIOR APPLICATION NUMBER: 60/115,435
; PRIOR FILING DATE: 1999-01-06
; PRIOR APPLICATION NUMBER: 60/118,848
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 106
```

```

; LENGTH: 2626
; TYPE: PRT
; ORGANISM: Streptomyces verticillius
; FEATURE:
; OTHER INFORMATION: ORF17
US-09-477-962-106
```

```

Query Match          61.4%: Score 35; DB 5; Length 2626;
Best Local Similarity 75.0%: Pred. No. 7.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 4 TAPKAPTE 11
    111: 111
Db 960 TAPRTPT 967
```

```

RESULT 11
US-09-758-469-419
; Sequence 419, Application US/09758469
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM008
; CURRENT APPLICATION NUMBER: US/09/758,469
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 419
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-469-419
```

```

Query Match          59.6%: Score 34; DB 5; Length 120;
Best Local Similarity 60.0%: Pred. No. 33;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 2 QQTAPKAPTE 11
    1: 111: 111
Db 55 QLSAPRAPSE 64
```

```

RESULT 12
PCT-US01-14827-14067
; Sequence 14067, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 14067
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(239)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-14827-14067
```

```

Query Match          59.6%: Score 34; DB 1; Length 239;
```

Best Local Similarity 66.7%; Pred. No. 72;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 QOQTAPKAPT 9  
|||  
Db 187 QDQTPKGP 195

RESULT 13  
US-09-758-470-399

; Sequence 399, Application US/09758470  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PM030  
; CURRENT APPLICATION NUMBER: US/09/758,470  
; CURRENT FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; NUMBER OF SEQ ID NOS: 722  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 399  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-758-470-399

Query Match 59.6%; Score 34; DB 5; Length 240;  
Best Local Similarity 70.0%; Pred. No. 72;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 QOQTAPKAPT 10  
|||  
Db 168 QOQCAPNAPT 177

RESULT 14  
US-09-760-462-27

; Sequence 27, Application US/09760462  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT256  
; CURRENT APPLICATION NUMBER: US/09/760,462  
; CURRENT FILING DATE: 2001-01-16  
; Prior Application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 27  
; LENGTH: 441  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (139)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (309)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (310)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-760-462-27

Query Match 59.6%; Score 34; DB 5; Length 441;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 QOQTAPKAPT 10

Db 20 QRRVAPSAPT 29  
|||  
PCT-US01-08656-7761

; Sequence 7761, Application PC/TUS0108656  
; GENERAL INFORMATION:  
; APPLICANT: Hysq, Inc  
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 21272-066  
; CURRENT APPLICATION NUMBER: PCT/US01/08656  
; CURRENT FILING DATE: 2001-04-16  
; PRIOR APPLICATION NUMBER: 09/522,929  
; PRIOR FILING DATE: 2000-04-18  
; PRIOR APPLICATION NUMBER: 09/770,160  
; PRIOR FILING DATE: 2001-01-26  
; NUMBER OF SEQ ID NOS: 10994  
; SOFTWARE: Custom  
; SEQ ID NO 7761  
; LENGTH: 143  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(143)  
; OTHER INFORMATION: Xaa = X or \* as defined in Table 2  
PCT-US01-08656-7761

Query Match 57.9%; Score 33; DB 1; Length 143;  
Best Local Similarity 100.0%; Pred. No. 60;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PKAPTE 11  
|||  
Db 111 PKAPTE 116

Search completed: August 15, 2001, 13:10:30  
Job time: 67 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 13:10:02 ; Search time 30.68 seconds  
(without alignments)  
8.054 Million cell updates/sec

Title: US-09-372-036-30  
Perfect score: 60  
Sequence: 1 STPVAPTOEVKK 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/p/odata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/p/odata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/p/odata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/p/odata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/p/odata/2/1aa/PCITUS.COMB.pep.\*  
6: /cgn2\_6/p/odata/2/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	12	2	US-08-456-670B-30
2	60	100.0	21	1	US-08-127-499A-27
3	60	100.0	21	1	US-08-482-847-27
4	60	100.0	232	2	US-08-456-670B-39
5	60	100.0	478	2	US-08-456-670B-40
6	60	100.0	484	1	US-08-127-499A-26
7	60	100.0	484	1	US-08-482-847-26
8	51	85.0	10	2	US-08-456-670B-31
9	41	68.3	9	2	US-08-456-670B-26
10	38	63.3	770	1	US-08-525-654A-1
11	38	63.3	771	1	US-08-525-654A-3
12	37	61.7	803	4	US-09-053-035-2
13	36	60.0	7	1	US-08-127-499A-28
14	36	60.0	7	1	US-08-482-847-28
15	36	60.0	1566	2	US-08-687-956A-23
16	35	58.3	12	2	US-08-389-011-16
17	35	58.3	12	2	US-08-403-917A-16
18	35	58.3	12	4	US-09-348-952A-16
19	35	58.3	26	2	US-08-389-011-7
20	35	58.3	26	3	US-08-403-917A-7
21	35	58.3	26	4	US-09-348-952A-7
22	35	58.3	33	2	US-08-244-951A-2
23	35	58.3	33	2	US-08-389-011-2
24	35	58.3	33	3	US-08-403-917A-2
25	35	58.3	33	4	US-09-348-952A-2
26	35	58.3	35	2	US-08-244-951A-6
27	35	58.3	106	3	US-08-776-404B-1

28	35	58.3	112	3	US-08-666-360-1	Sequence 1, Appl
29	35	58.3	126	6	5514582-35	Patent No. 5514582
30	35	58.3	250	4	US-09-216-295-14	Sequence 14, Appl
31	35	58.3	351	1	US-08-159-969-2	Sequence 2, Appl
32	35	58.3	352	2	US-08-726-306A-17	Sequence 17, Appl
33	35	58.3	391	2	US-08-244-951A-10	Sequence 10, Appl
34	35	58.3	391	2	US-08-389-011-23	Sequence 23, Appl
35	35	58.3	391	3	US-08-403-917A-23	Sequence 23, Appl
36	35	58.3	391	4	US-09-348-952A-23	Sequence 23, Appl
37	35	58.3	441	4	US-08-244-603A-1	Sequence 1, Appl
38	35	58.3	605	2	US-08-687-956A-1	Sequence 1, Appl
39	35	58.3	1466	6	5256642-6	Patent No. 5256642
40	35	58.3	1466	6	5472939-6	Patent No. 5472939
41	35	58.3	1537	6	5256642-5	Patent No. 5256642
42	35	58.3	1537	6	5472939-5	Patent No. 5472939
43	35	58.3	1847	6	5256642-10	Patent No. 5256642
44	35	58.3	1847	6	5472939-10	Patent No. 5472939
45	35	58.3	2039	6	5256642-2	Patent No. 5256642

## ALIGNMENTS

RESULT 1  
US-08-456-670B-30  
Sequence 30 Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELTER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 169401  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-436-670B-30

Query Match 100.0%; Score 60; DB 1; Length 12;  
Best Local Similarity 100.0%; Pred. No. 8.7e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STVPAPTOEVKK 12  
DB 1 STVPAPTOEVKK 12

RESULT 2  
US-08-127-499A-27  
Sequence 27, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-27

Query Match 100.0%; Score 60; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STVPAPTOEVKK 12  
DB 3 STVPAPTOEVKK 14

RESULT 3  
US-08-482-847-27  
Sequence 27, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-27

Query Match 100.0%; Score 60; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STVPAPTOEVKK 12  
DB 3 STVPAPTOEVKK 14

RESULT 4  
US-08-456-670B-39  
Sequence 39, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKWIELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING

;; TITLE OF INVENTION: LISTERIAS  
;; NUMBER OF SEQUENCES: 43  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
;; STREET: 2200 CLARENDON BLVD., SUITE 1400  
;; CITY: ARLINGTON  
;; STATE: VIRGINIA  
;; COUNTRY: US  
;; ZIP: 22201  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,670B  
;; FILING DATE: 01-JUN-1995  
;; CLASSIFICATION: 435  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/412,227  
;; FILING DATE: 27-MAR-1995  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/075,248  
;; FILING DATE: 11-JUN-1993  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: DE 4239567.4  
;; FILING DATE: 25-NOV-1992  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: DE 4219111.4  
;; FILING DATE: 11-JUN-1992  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: HAMLET-KING, DIANA  
;; REGISTRATION NUMBER: 33,302  
;; REFERENCE/DOCKET NUMBER: MERCK 1694D1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-243-6333  
;; TELEFAX: 703-243-6410  
;;  
;; TELEX: 64191  
;;  
;; INFORMATION FOR SEQ ID NO: 39:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 232 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; ORIGINAL SOURCE:  
;; ORGANISM: Listeria monocytogenes  
;; STRAIN: EGD  
;;  
;; US-08-456-670B-39

Query Match 100.0%; Score 60; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 0.0022;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STPVAPTOEVKK 12  
DB 77 STPVAPTOEVKK 88

RESULT 5  
US-08-456-670B-40  
;; Sequence 40, Application US/08456670B  
;; Patent No. 5932415  
;; GENERAL INFORMATION:  
;; APPLICANT: SCHUBERT, PETER  
;; APPLICANT: NEUMANN, SIEGFRIED  
;; APPLICANT: PAMELZIK, MARTINA  
;; APPLICANT: LINKEWILLER, WINFRIED  
;; APPLICANT: BURGER, CHRISTA  
;; APPLICANT: HOFMANN, GOTTFRIED  
;; APPLICANT: BUBERT, ANDREAS  
;; APPLICANT: GOEBEL, WERNER  
;; APPLICANT: KOHLER, STEFAN

;; TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
;; TITLE OF INVENTION: LISTERIAS  
;; NUMBER OF SEQUENCES: 43  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
;; STREET: 2200 CLARENDON BLVD., SUITE 1400  
;; CITY: ARLINGTON  
;; STATE: VIRGINIA  
;; COUNTRY: US  
;; ZIP: 22201  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/456,670B  
;; FILING DATE: 01-JUN-1995  
;; CLASSIFICATION: 435  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/412,227  
;; FILING DATE: 27-MAR-1995  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/075,248  
;; FILING DATE: 11-JUN-1993  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: DE 4239567.4  
;; FILING DATE: 25-NOV-1992  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: DE 4219111.4  
;; FILING DATE: 11-JUN-1992  
;;  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: HAMLET-KING, DIANA  
;; REGISTRATION NUMBER: 33,302  
;; REFERENCE/DOCKET NUMBER: MERCK 1694D1  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-243-6333  
;; TELEFAX: 703-243-6410  
;;  
;; TELEX: 64191  
;;  
;; INFORMATION FOR SEQ ID NO: 40:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 478 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; ORIGINAL SOURCE:  
;; ORGANISM: Listeria monocytogenes  
;; STRAIN: EGD  
;;  
;; US-08-456-670B-40

Query Match 100.0%; Score 60; DB 2; Length 478;  
Best Local Similarity 100.0%; Pred. No. 0.0047;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STPVAPTOEVKK 12  
DB 146 STPVAPTOEVKK 157

RESULT 6  
US-08-127-499A-26  
;; Sequence 26, Application US/08127499A  
;; Patent No. 5510264  
;; GENERAL INFORMATION:  
;; APPLICANT: VAN ALSTYNE, Diane  
;; APPLICANT: SHARMA, Lawrence Rajendra  
;; TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
;; TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
;; NUMBER OF SEQUENCES: 40  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Foley & Lardner  
;; STREET: 3000 K Street, N.W., Suite 500

CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-26

Query Match 100.0%; Score 60; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVAPTOEVKK 12  
Db 146 STPVAPTOEVKK 157

RESULT 7  
US-08-482-847-26  
Sequence 26, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
NUMBER OF SEQUENCES: 40  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-26

Query Match 100.0%; Score 60; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 0.0048;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPVAPTOEVKK 12  
Db 146 STPVAPTOEVKK 157

RESULT 8  
US-08-456-670B-31  
Sequence 31, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, Siegfried  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKEWELER, WILFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLER, WHITE, ZELANO & BRANTIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410



TELEX: 64191  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-31

Query Match 85.0%; Score 51; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 PVAPTOEVKK 12  
|||||  
Db 1 PVAPTOEVKK 10

RESULT 9  
US-08-456-670B-26  
Sequence 26, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:

APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, STEFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKWEILER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BOBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:

ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456, 670B

FILING DATE: 01-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/412,227

FILING DATE: 27-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/075,248

FILING DATE: 11-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4239567.4

FILING DATE: 25-NOV-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 421911.4

FILING DATE: 11-JUN-1992

ATTORNEY/AGENT INFORMATION:

NAME: HAMLET-KING, DIANA

REGISTRATION NUMBER: 33,302

REFERENCE/DOCKET NUMBER: MERCK 1694D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-243-6333

TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-26

Query Match 68.3%; Score 41; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.3e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 STPVAPTO 8  
|||||  
Db 2 STPVAPTO 9

RESULT 10  
US-08-525-654A-1  
Sequence 1, Application US/08525654A  
Patent No. 5736356  
GENERAL INFORMATION:

APPLICANT: SANO, KOICHIRO  
APPLICANT: KUMAZAWA, YOSHIYUKI  
APPLICANT: YASEUDA, HISASHI  
APPLICANT: SEGURO, KATSUYA  
APPLICANT: MOTOKI, MASAO  
TITLE OF INVENTION: TRANSGLUTAMINASE ORIGINATED FROM  
NUMBER OF SEQUENCES: 150  
CORRESPONDENCE ADDRESS:

ADDRESSEE: P.C. SPIVAK, MCCLELLAND, MATER & NEUSTADT,

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

CITY: ARLINGTON

STATE: VA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/525, 654A

FILING DATE: 28-SEP-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6/8283

FILING DATE: 28-JAN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7/3876

FILING DATE: 13-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 10-760-0 PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 770 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORIGINAL SOURCE:  
ORGANISM: Crassostrea gigas  
US-08-525-654A-1

Query Match 63.3%; Score 38; DB 1; Length 770;  
Best Local Similarity 58.3%; Pred. No. 58;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 STPVAPTOEVKK 12  
|:|:| |  
Db 711 SSPKPGDEVKK 722

RESULT 11  
US-08-525-654A-3  
Sequence 3, Application US/08525654A  
Patent No. 5736356  
GENERAL INFORMATION:  
APPLICANT: SANO, KOICHIRO  
APPLICANT: KUMAZAWA, YOSHIYUKI  
APPLICANT: YASEUDA, HISASHI  
APPLICANT: SEGURO, KATSUYA  
APPLICANT: MOTOKI, MASAO  
TITLE OF INVENTION: TRANSGUTAMINASE ORIGINATED FROM  
TITLE OF INVENTION: CRASSOSTREA GIGAS  
NUMBER OF SEQUENCES: 150  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/525,654A  
FILING DATE: 28-SEP-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6/8283  
FILING DATE: 28-JAN-1994  
PRIOR APPLICATION DATA: JP 7/3876  
APPLICATION NUMBER: 24,618  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 10-760-0 PCT  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 771 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Crassostrea gigas  
US-08-525-654A-3

Query Match 63.3%; Score 38; DB 1; Length 771;  
Best Local Similarity 58.3%; Pred. No. 58;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 STPVAPTOEVKK 12  
|:|:| |  
Db 712 SSPKPGDEVKK 723

RESULT 12  
US-09-063-035-2  
Sequence 2, Application US/09063035  
Patent No. 6160091  
GENERAL INFORMATION:  
APPLICANT: PEUKERT, Karen; HAENEL, Frank; and ELLERS,  
APPLICANT: Martin  
TITLE OF INVENTION: Myc-binding zinc finger proteins,  
TITLE OF INVENTION: their preparation and their use  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Keil & Weinlauf  
STREET: 1101 Connecticut Avenue  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.2 MB storage  
COMPUTER: IBM AT-compatible, 80486 processor  
OPERATING SYSTEM: MS-DOS version 6.1  
SOFTWARE: Wordperfect version 8.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/063,035  
FILING DATE: 21-APR-1998  
CLASSIFICATION: 514  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 803 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-063-035-2

Query Match 61.7%; Score 37; DB 4; Length 803;  
Best Local Similarity 41.7%; Pred. No. 91;  
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 STPVAPTOEVKK 12  
|:|:| |  
Db 156 STPGPSRDLKE 167

RESULT 13  
US-08-127-499A-28  
Sequence 28, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993

ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-28

Query Match 60.0%; Score 36; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 PTOEVKK 12  
|||  
Db 1 PTOEVKK 7

RESULT 14  
US-08-482-847-28  
Sequence 28, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHAWA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THERETO, AND USES THEREOF  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-28

Query Match 60.0%; Score 36; DB 1; Length 7;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 PTOEVKK 12  
|||  
Db 1 PTOEVKK 7

RESULT 15  
US-08-687-956A-23  
Sequence 23, Application US/08687956A  
Patent No. 5861157  
GENERAL INFORMATION:  
APPLICANT: BURNIE, JAMES P  
APPLICANT: MATTHEWS, RUTH C  
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF  
INFECTIONS DUE TO STREPTOCOCCI AND ENTEROCOCCI  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PILLSBURY, MADISON & SUTRO, LLP  
STREET: 1100 NEW YORK AVENUE, N.W., EAST TOWER, 9TH  
FLOOR  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,956A  
FILING DATE: 29-JUL-1996  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9401689.6  
FILING DATE: 28-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: 50885/222892  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202/861-3000  
TELEFAX: 202/822-0944  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1566 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus sobrinus  
US-08-687-956A-23

Query Match 60.0%; Score 36; DB 2; Length 1566;  
Best Local Similarity 70.0%; Pred. No. 2.8e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 3 PVAPTQEVKK 12  
|||  
Db 887 PVEPTVEVKK 896

Search completed: August 15, 2001, 13:10:02  
Job time: 39 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 13:09:23 ; Search time 30.68 Seconds  
(without alignments)  
6.040 Million cell updates/sec

Title: US-09-372-036-26

Perfect score: 45

Sequence: 1 VSTRPVAPQTQ 9

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCYUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45	100.0	9	2	US-08-456-670B-26
2	45	100.0	21	1	US-08-127-499A-27
3	45	100.0	21	1	US-08-482-847-27
4	45	100.0	232	2	US-08-456-670B-39
5	45	100.0	478	2	US-08-456-670B-39
6	45	100.0	484	1	US-08-127-499A-26
7	45	100.0	484	1	US-08-482-847-26
8	41	91.1	12	2	US-08-456-670B-30
9	32	71.1	10	2	US-08-456-670B-31
10	32	71.1	20	2	US-08-456-670B-17
11	32	71.1	493	3	US-09-090-808-4
12	31	68.9	45	3	US-08-856-074A-29
13	31	68.9	250	4	US-09-216-295-14
14	31	68.9	469	2	US-08-484-126-1
15	31	68.9	763	2	US-08-473-123-2
16	31	68.9	763	2	US-08-472-478-2
17	31	68.9	763	2	US-08-463-081B-8
18	31	68.9	763	2	US-08-461-379A-8
19	31	68.9	763	2	US-08-462-390B-8
20	31	68.9	763	3	US-08-463-074B-8
21	31	68.9	763	3	US-08-465-585C-8
22	31	68.9	763	3	US-08-652-446-8
23	31	68.9	2476	2	US-08-276-967-2
24	30	66.7	282	1	US-07-712-476A-5
25	30	66.7	296	1	US-07-712-476A-1
26	30	66.7	341	1	US-08-356-180-4
27	30	66.7	404	1	US-08-696-770-2

28	30	66.7	404	2	US-09-015-557-2	Sequence 2, Appl1
29	30	66.7	433	1	US-07-672-483-4	Sequence 4, Appl1
30	30	66.7	604	2	US-08-468-576B-12	Sequence 12, Appl1
31	30	66.7	604	2	US-08-468-576B-12	Sequence 12, Appl1
32	30	66.7	604	3	US-08-468-576B-12	Sequence 12, Appl1
33	30	66.7	1308	2	US-08-996-644-2	Sequence 2, Appl1
34	30	66.7	1308	3	US-09-352-552-2	Sequence 2, Appl1
35	30	66.7	1436	2	US-08-652-971-2	Sequence 2, Appl1
36	30	66.7	1436	2	US-08-991-258A-2	Sequence 2, Appl1
37	30	66.7	1436	2	US-08-769-399-2	Sequence 2, Appl1
38	30	66.7	1436	2	US-08-991-953A-2	Sequence 2, Appl1
39	30	66.7	1618	1	US-07-853-913-4	Sequence 4, Appl1
40	30	66.7	15281	2	US-08-471-119A-2	Sequence 2, Appl1
41	29	64.4	15	1	US-08-116-733-15	Sequence 15, Appl1
42	29	64.4	310	1	US-08-078-683A-3	Sequence 3, Appl1
43	29	64.4	310	2	US-08-488-199-4	Sequence 3, Appl1
44	29	64.4	327	2	US-08-651-818A-3	Sequence 3, Appl1
45	29	64.4	327	4	US-09-184-826-3	Sequence 3, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-456-670B-26  
Sequence 26, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIGFRIED  
APPLICANT: PAMELLEIK, MARTINA  
APPLICANT: LINWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 421911.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAWLEY-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: *Listeria monocytogenes*  
STRAIN: EGD  
US-08-456-6708-26

Query Match 100.0%; Score 45; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPTO 9  
Db 1 VSTPVAPTO 9

RESULT 2  
US-08-127-499A-27  
Sequence 27, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-27

Query Match 100.0%; Score 45; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.056;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPTO 9  
Db 2 VSTPVAPTO 10

RESULT 3  
US-08-482-847-27  
Sequence 27, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-27

Query Match 100.0%; Score 45; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.056;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPTO 9  
Db 2 VSTPVAPTO 10

RESULT 4  
US-08-456-6708-39  
Sequence 39, Application US/084566708  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, Siegfried  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINWELTER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING

```

: TITLE OF INVENTION: LISTERIAS
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
: STREET: 2200 CLARENDON BLVD., SUITE 1400
: CITY: ARLINGTON
: STATE: VIRGINIA
: COUNTRY: US
: ZIP: 22201
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/456,670B
: FILING DATE: 01-JUN-1995
: CLASSIFICATION: 435
: APPLICATION NUMBER: US 08/412,227
: FILING DATE: 27-MAR-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/075,248
: FILING DATE: 11-JUN-1993
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: DE 4239567.4
: FILING DATE: 25-NOV-1992
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: DE 4219111.4
: FILING DATE: 11-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: HAMLET-KING, DIANA
: REGISTRATION NUMBER: 33,302
: REFERENCE/DOCKET NUMBER: MERCK 1694D1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-243-6333
: TELEFAX: 703-243-6410
: TELEX: 64191
: INFORMATION FOR SEQ ID NO: 39:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 232 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: ORIGINAL SOURCE:
: ORGANISM: Listeria monocytogenes
: STRAIN: EGD
: US-08-456-670B-39

Query Match 100.0%; Score 45; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPTQ 9
Db 76 VSTPVAPTQ 84

RESULT 5
US-08-456-670B-40
: Sequence 40, Application US/08456670B
: Patent No. 5932415
: GENERAL INFORMATION:
: APPLICANT: SCHUBERT, PETER
: APPLICANT: NEUMANN, SIEGFRIED
: APPLICANT: PAMELLEK, MARTINA
: APPLICANT: LINXWEILER, WINFRIED
: APPLICANT: BURGER, CHRISTA
: APPLICANT: HOFMANN, GOTTFRIED
: APPLICANT: BOBERT, ANDREAS
: APPLICANT: GOEBEL, WERNER
: APPLICANT: KOHLER, STEFAN
```

```

: TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
: STREET: 2200 CLARENDON BLVD., SUITE 1400
: CITY: ARLINGTON
: STATE: VIRGINIA
: COUNTRY: US
: ZIP: 22201
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/456,670B
: FILING DATE: 01-JUN-1995
: CLASSIFICATION: 435
: APPLICATION NUMBER: US 08/412,227
: FILING DATE: 27-MAR-1995
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/075,248
: FILING DATE: 11-JUN-1993
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: DE 4239567.4
: FILING DATE: 25-NOV-1992
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: DE 4219111.4
: FILING DATE: 11-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: HAMLET-KING, DIANA
: REGISTRATION NUMBER: 33,302
: REFERENCE/DOCKET NUMBER: MERCK 1694D1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 703-243-6333
: TELEFAX: 703-243-6410
: TELEX: 64191
: INFORMATION FOR SEQ ID NO: 40:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 478 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: ORIGINAL SOURCE:
: ORGANISM: Listeria monocytogenes
: STRAIN: EGD
: US-08-456-670B-40

Query Match 100.0%; Score 45; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPTQ 9
Db 145 VSTPVAPTQ 153

RESULT 6
US-08-127-499A-26
: Sequence 26, Application US/08127499A
: Patent No. 5510264
: GENERAL INFORMATION:
: APPLICANT: VAN ALSTYNE, Diane
: APPLICANT: SHARMA, Lawrence Rajendra
: TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
: TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES
: NUMBER OF SEQUENCES: 40
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
```

CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-26

Query Match 100.0%; Score 45; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPTO 9  
|||||||  
DB 145 VSTPVAPTO 153

RESULT 7  
US-08-482-847-26  
Sequence 26, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-26

Query Match 100.0%; Score 45; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPTO 9  
|||||||  
DB 145 VSTPVAPTO 153

RESULT 8  
US-08-456-670B-30  
Sequence 30, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, STEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINKWELTER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410



TELEX: 64191  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-30

Query Match 91.1%; Score 41; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVAPQ 9  
|||||  
DB 1 STPVAPQ 8

RESULT 9  
US-08-456-670B-31  
Sequence 31, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, STEFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BOBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567,4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111,4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333

TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-31

Query Match 71.1%; Score 32; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPQ 9  
|||||  
DB 1 PVAPQ 6

RESULT 10  
US-08-456-670B-17  
Sequence 17, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, STEFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINWELER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BOBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567,4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111,4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333

TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: listeria monocytogenes  
STRAIN: EGD  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1-7  
OTHER INFORMATION: /note= "SOME OR ALL Xaa AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 14-20  
OTHER INFORMATION: /note= "SOME OR ALL Xaa AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
US-08-456-670B-17

Query Match 71.1%; Score 32; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPTO 9  
|||||  
DB 8 PVAPTO 13

RESULT 11  
US-09-090-808-4  
Sequence 4, Application US/09090808  
Patent No. 6127159  
GENERAL INFORMATION:  
APPLICANT: Fuller, Margaret  
TITLE OF INVENTION: Mitofusin Genes and their  
TITLE OF INVENTION: Uses  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bozicevic & Reed, LLP  
STREET: 285 Hamilton Avenue, Suite 200  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/090,808  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/048,961  
FILING DATE: 06-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J  
REGISTRATION NUMBER: 36,677  
REFERENCE/DOCKET NUMBER: SUN-63P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-327-3400  
TELEFAX: 650-327-3231  
TELEX:  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:

LENGTH: 493 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-090-808-4

Query Match 71.1%; Score 32; DB 3; Length 493;  
Best Local Similarity 85.7%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 STPVAPT 8  
|||||  
DB 313 STPVAPT 319

RESULT 12  
US-08-856-074A-29

Sequence 29, Application US/08856074A  
Patent No. 6004798

GENERAL INFORMATION:

APPLICANT: Anderson, W. French

APPLICANT: Wu, Bonnie W.

TITLE OF INVENTION: Retroviral Envelopes Having

TITLE OF INVENTION: Modified Hypervariable polyprolline Regions

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: Carrella, Byrne, Bain, Gillfillan, Cecchi, Stewart & Olstein

STREET: 6 Becker Farm Road

CITY: Roseland

STATE: New Jersey

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch diskette

COMPUTER: IBM PS/2

OPERATING SYSTEM: WINDOWS 95

SOFTWARE: Word 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/856,074A

FILING DATE: 14-May-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Olstein, Elliot M.

REGISTRATION NUMBER: 24,025

REFERENCE/DOCKET NUMBER: 271010-378

TELECOMMUNICATION INFORMATION:

TELEPHONE: 973-994-1700

TELEFAX: 973-994-1744

INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:

LENGTH: 45 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: polypeptide

FEATURE:

NAME/KEY: Ecotropic hypervariable polyprolline region of a retroviral envelope

US-08-856-074A-29

Query Match 68.9%; Score 31; DB 3; Length 45;  
Best Local Similarity 71.4%; Pred. No. 31;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPTQ 9  
|||||  
DB 36 TPVAPTQ 42

RESULT 13  
US-09-216-295-14  
; Sequence 14, Application US/09216295  
; Patent No. 6288328  
; GENERAL INFORMATION:  
; APPLICANT: Mitchinson, Colln  
; TITLE OF INVENTION: No. 6288328el Variant Egit-Like Cellulase Compositions  
; FILE REFERENCE: GC555  
; CURRENT APPLICATION NUMBER: US/09/216,295  
; CURRENT FILING DATE: 1998-12-18  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Fusarium javanicum (2)  
US-09-216-295-14

Query Match 68.9%; Score 31; DB 4; Length 250;  
Best Local Similarity 55.6%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VSTPVPQTQ 9  
|:|:|:|:  
DB 16 VAPTPTTPE 24

RESULT 14  
US-08-484-126-1  
; Sequence 1, Application US/08484126  
; Patent No. 5985655  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, W. French  
; APPLICANT: Baltuckl, Leon F.  
; APPLICANT: Mason, James M.  
; TITLE OF INVENTION: Targetable Vector Particles  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Carella, Byrne, Baln, Gilfillan,  
; ADDRESSEE: Cecchi, Stewart & Olstein  
; STREET: 6 Becker Farm Road  
; CITY: Roseland  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Word Perfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,126  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/326,347  
; FILING DATE: 20-OCT-1994  
; APPLICATION NUMBER: 08/973,307  
; FILING DATE: 09-NOV-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lillie, Raymond J.  
; REGISTRATION NUMBER: 31,778  
; REFERENCE/DOCKET NUMBER: 271010-281  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-994-1700  
; TELEFAX: 201-994-1744  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 469 amino acids  
; TYPE: amino acid

STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Ecotropic gp70 Protein  
US-08-484-126-1

Query Match 68.9%; Score 31; DB 2; Length 469;  
Best Local Similarity 71.4%; Pred. No. 3.7e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVPQTQ 9  
|:|:|:|:  
DB 298 TPSPPTQ 304

RESULT 15  
US-08-473-122-2  
; Sequence 2, Application US/08473122  
; Patent No. 5652340  
; GENERAL INFORMATION:  
; APPLICANT: KOHWI-SHIGEMATSU, TERUMI  
; APPLICANT: KOHWI, YOSHINORI  
; APPLICANT: DICKINSON, LILIANE A  
; TITLE OF INVENTION: MATRIX-ASSOCIATING DNA-BINDING PROTEIN,  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING THE SAME AND METHODS FOR DETECTING  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CAMPBELL AND FLORES  
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700  
; CITY: SAN DIEGO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/473,122  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/934,034  
; FILING DATE: 21-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CAMPBELL, CATHRYN  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LJ 9293  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619 535-9001  
; TELEFAX: 619 535-8949  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 763 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-473-122-2

Query Match 68.9%; Score 31; DB 1; Length 763;  
Best Local Similarity 71.4%; Pred. No. 6.2e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVP 7  
|:|:|:|:  
DB 308 VSTPISP 314

Search completed: August 15, 2001, 13:10:02  
Job time: 39 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:32:55 ; Search time 56.84 Seconds  
(without alignments)  
9.599 Million cell updates/sec

Title: US-09-372-036-26

Perfect score: 9

Sequence: 1 VSTRPVAPQTQ 9

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

A.Geneseq\_0601.\*  
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100.0	9	14	AA845168	Listeria p60 pepti
2	100.0	21	16	AA873891	Listeria monocytog
3	100.0	478	14	AA845178	Listeria p60 prote
4	100.0	484	16	AA873913	Listeria monocytog
5	88.9	12	14	AA845172	Listeria p60 pepti
6	66.7	6	14	AA845615	Listeria p60 pepti
7	66.7	10	14	AA845173	Listeria p60 pepti
8	66.7	20	14	AA845159	Listeria p60 pepti
9	66.7	157	20	AA873718	Amino acid sequenc
10	66.7	343	21	AA841666	Human ORFX ORF1430
11	66.7	750	20	AA805477	C. albicans Rbtl p

12	66.7	999	22	AA859817	TuTd protein #8.
13	66.7	1592	22	AA859827	Protein #4 encoded
14	55.6	7	19	AA821365	Human HUPF-I mutan
15	55.6	9	19	AA870120	Peptide produced b
16	55.6	11	19	AA870126	Peptide produced b
17	55.6	11	22	AA872848	Human p53 A76T mut
18	55.6	20	20	AA882295	Dynamn 41.20 Zela
19	55.6	23	20	AA882310	Transport peptide
20	55.6	25	16	AA879882	Signal transductio
21	55.6	27	20	AA878130	Human secreted pro
22	55.6	43	20	AA882294	Dynamn 41.2 Zelan
23	55.6	45	20	AA801143	Secreted protein e
24	55.6	48	21	AA833087	Plus radiata tran
25	55.6	57	21	AA844975	Human secreted pro
26	55.6	59	21	AA844976	Human secreted pro
27	55.6	72	21	AA839069	Human secreted pro
28	55.6	87	20	AA827568	Human secreted pro
29	55.6	89	21	AA841518	Human ORFX ORF1282
30	55.6	94	15	AA847117	Complete sequence
31	55.6	98	21	AA854394	Arabidopsis thalia
32	55.6	101	21	AA82898	CUB domain from BM
33	55.6	102	22	AA872875	Human p53 A76T/V12
34	55.6	104	19	AA856088	Murine monocyte ch
35	55.6	104	19	AA857322	Mouse monocyte che
36	55.6	112	21	AA854393	Arabidopsis thalia
37	55.6	114	19	AA811113	H. pylori ORF hp6e
38	55.6	114	21	AA803726	Human secreted pro
39	55.6	121	21	AA854455	Zea mays protein f
40	55.6	125	21	AA840264	Human ORFX ORF28 p
41	55.6	125	21	AA854454	Zea mays protein f
42	55.6	130	19	AA877657	Staphylococcus aur
43	55.6	130	21	AA853631	Human colon cancer
44	55.6	132	21	AA855449	Human prostate can
45	55.6	151	21	AA845699	Human cancer assoc
46	55.6	152	21	AA854453	Zea mays protein f
47	55.6	152	22	AA879973	Corynebacterium g1
48	55.6	152	22	AA879981	Corynebacterium g1
49	55.6	152	22	AA879987	Corynebacterium g1
50	55.6	152	22	AA879987	Plant SDF encoded
51	55.6	158	21	AA824690	Plant SDF encoded
52	55.6	158	21	AA825075	Plant SDF encoded
53	55.6	160	21	AA812578	Zea mays protein f
54	55.6	160	21	AA810791	M. avium Innazine
55	55.6	168	22	AA879120	Corynebacterium g1
56	55.6	168	22	AA879813	Corynebacterium g1
57	55.6	179	21	AA855986	Arabidopsis thalia
58	55.6	179	21	AA855942	Arabidopsis thalia
59	55.6	180	21	AA855985	Arabidopsis thalia
60	55.6	181	21	AA855985	Arabidopsis thalia
61	55.6	181	21	AA855985	Arabidopsis thalia
62	55.6	181	21	AA855985	Arabidopsis thalia
63	55.6	184	22	AA880053	Arabidopsis thalia
64	55.6	185	21	AA854822	Corynebacterium g1
65	55.6	191	20	AA894495	Arabidopsis thalia
66	55.6	191	16	AA869639	Human delta-2 prot
67	55.6	192	17	AA889510	Hepatitis C virus
68	55.6	193	14	AA833986	Hepatitis C virus
69	55.6	193	21	AA806865	HCT27 El protein.
70	55.6	196	21	AA810819	Arabidopsis thalia
71	55.6	196	21	AA855984	Arabidopsis thalia
72	55.6	202	18	AA859940	Arabidopsis thalia
73	55.6	202	18	AA859940	BRCA2 cancer suscep
74	55.6	202	21	AA809042	Arabidopsis thalia
75	55.6	202	21	AA838651	Arabidopsis thalia
76	55.6	205	21	AA806864	Arabidopsis thalia
77	55.6	205	21	AA810818	Arabidopsis thalia
78	55.6	214	21	AA850386	Human uncoupling p
79	55.6	216	22	AA875329	Human secreted pro
80	55.6	217	21	AA804638	Arabidopsis thalia
81	55.6	219	20	AA830653	Human secreted pro
82	55.6	226	21	AA832101	Human secreted pro
83	55.6	238	21	AA813699	Chlamydia sp. prot
84	55.6	243	19	AA880702	S. pneumoniae cati
		247	20	AA827338	Group B Streptococ

85	5	55.6	253	21	AA05005	Arabidopsis thalia	158	5	55.6	500	20	AAW94496	Human delta-2 prot
86	5	55.6	255	20	AAV37385	Amino acid sequenc	159	5	55.6	501	21	AAAB32978	Pinus radiata tran
87	5	55.6	256	20	AA04637	Arabidopsis thalia	160	5	55.6	503	21	AAV68936	Amino acids 29-531
88	5	55.6	264	13	AA029638	PCPD ORF 3. Chlam	161	5	55.6	504	21	AAV68935	Amino acids 28-531
89	5	55.6	276	20	AAAB1577	EDG-1-like G-prote	162	5	55.6	505	21	AAV68918	Amino acids 28-532
90	5	55.6	280	21	AA04636	Arabidopsis thalia	163	5	55.6	505	21	AAV68934	Amino acids 27-531
91	5	55.6	281	21	AA05004	Arabidopsis thalia	164	5	55.6	506	21	AAV68917	Amino acids 27-532
92	5	55.6	283	21	AAAB12148	Hydrophobic domain	165	5	55.6	506	21	AAV68933	Amino acids 26-531
93	5	55.6	283	21	AAV94952	Human secreted pro	166	5	55.6	507	21	AAV68916	Amino acids 26-532
94	5	55.6	286	17	AAW03566	Mycobacterium tube	167	5	55.6	507	21	AAV68932	Amino acids 25-531
95	5	55.6	302	21	AA05003	Arabidopsis thalia	168	5	55.6	508	21	AAV68915	Amino acids 25-532
96	5	55.6	305	17	AAW0650	Protein phosphatas	169	5	55.6	508	21	AAV68931	Amino acids 24-531
97	5	55.6	323	21	AAV75550	Neisseria gonorrhe	170	5	55.6	509	19	AAW80935	Amino acid sequenc
98	5	55.6	323	22	AAAB51240	TPAI protein sequ	171	5	55.6	510	21	AAV68914	Amino acids 24-532
99	5	55.6	325	17	AAW03565	Mycobacterium tube	172	5	55.6	524	21	AAV68913	Amino acids 23-532
100	5	55.6	327	21	AAV97281	Fibronectin attach	173	5	55.6	525	21	AAV68942	Amino acids 29-552
101	5	55.6	327	21	AA099947	Arabidopsis thalia	174	5	55.6	525	21	AAV68941	Amino acids 28-552
102	5	55.6	328	21	AA099946	Arabidopsis thalia	175	5	55.6	526	21	AAV68924	Amino acids 28-553
103	5	55.6	332	18	AAW32418	Mycobacterium tube	176	5	55.6	526	21	AAV68940	Amino acids 27-552
104	5	55.6	332	18	AAW32350	M. tuberculosis Im	177	5	55.6	527	21	AAV68923	Amino acids 27-553
105	5	55.6	332	19	AAW81683	Mycobacterium tube	178	5	55.6	527	21	AAV68939	Amino acids 26-552
106	5	55.6	332	19	AAW64322	M. tuberculosis an	179	5	55.6	528	21	AAV68922	Amino acids 26-553
107	5	55.6	332	20	AAV39083	M. tuberculosis re	180	5	55.6	528	21	AAV68938	Amino acids 25-552
108	5	55.6	332	20	AAV38945	Arabidopsis thalia	181	5	55.6	529	21	AAV68921	Amino acids 25-553
109	5	55.6	336	21	AA046021	Arabidopsis thalia	182	5	55.6	529	21	AAV68937	Amino acids 24-552
110	5	55.6	343	21	AA046020	Arabidopsis thalia	183	5	55.6	530	21	AAV68920	Amino acids 24-553
111	5	55.6	351	21	AA046019	Arabidopsis thalia	184	5	55.6	531	21	AAV68919	Amino acids 23-553
112	5	55.6	357	21	AAV75397	Neisseria gonorrhe	185	5	55.6	535	22	AAAB49707	Small round struct
113	5	55.6	359	16	AAV72715	hisc gene product	186	5	55.6	543	20	AAW85610	Secreted protein c
114	5	55.6	370	13	AA026183	MS2-pgpD protein.	187	5	55.6	559	19	AAW50909	Alteromonas fortis
115	5	55.6	371	20	AAV29193	Amino acid sequenc	188	5	55.6	561	19	AAW63701	Human HSK1 protein
116	5	55.6	375	21	AA099945	Arabidopsis thalia	189	5	55.6	566	20	AAV49068	Amino acid sequenc
117	5	55.6	377	20	AAV24072	Salmonella typhimu	190	5	55.6	566	22	AAAB31932	Pseudomonas fluore
118	5	55.6	382	19	AAW44865	Human TP33 telomer	191	5	55.6	583	20	AAV55919	Human secreted pro
119	5	55.6	384	20	AAV06411	Human EDG-7 recept	192	5	55.6	588	20	AAW78199	Endocarditis speci
120	5	55.6	384	20	AAV06412	Human EDG-7 recept	193	5	55.6	605	16	AAV79625	Human delta-2 matu
121	5	55.6	384	21	AAV69500	A human G-protein	194	5	55.6	612	17	AAW02133	Lipoma preferred p
122	5	55.6	384	21	AAW90862	Human edg6 protein	195	5	55.6	612	19	AAW85816	H. pylori GHP 130
123	5	55.6	387	22	AAAB1911	R. anatisestifer o	196	5	55.6	651	19	AAV68935	Gravlin polypeptide
124	5	55.6	396	19	AAW53868	Gravlin polypeptide	197	5	55.6	651	21	AAAB18410	Human gravlin PKA r
125	5	55.6	396	21	AAAB18408	Human gravlin PKA r	198	5	55.6	652	20	AAV39225	M. tuberculosis fu
126	5	55.6	402	20	AAV24475	Mouse p14. Mus sp	199	5	55.6	652	20	AAV39082	M. tuberculosis fus
127	5	55.6	405	10	AAW90458	Snake venom fibrol	200	5	55.6	655	14	AAAB31041	snrk polypeptide.
128	5	55.6	405	11	AAW04082	Gene product of so	201	5	55.6	659	20	AAW94497	Human delta-2 matu
129	5	55.6	410	21	AAW18329	Arabidopsis thalia	202	5	55.6	659	21	AAV68930	Amino acids 28-666
130	5	55.6	411	10	AAW90457	Snake venom fibrol	203	5	55.6	660	21	AAV68929	Amino acids 27-666
131	5	55.6	411	11	AAW04081	Gene product of so	204	5	55.6	661	21	AAV68928	Amino acids 26-666
132	5	55.6	421	11	AA018328	Arabidopsis thalia	205	5	55.6	662	20	AAV31940	Human rad17 cell c
133	5	55.6	422	18	AAW62528	Branched chain ket	206	5	55.6	662	21	AAV68927	Amino acids 25-666
134	5	55.6	422	22	AAAB1181	Pseudomonas putida	207	5	55.6	663	20	AAV31941	Human rad17 cell c
135	5	55.6	422	22	AAW66454	Protein encoded by	208	5	55.6	663	21	AAV68926	Amino acids 24-666
136	5	55.6	428	20	AAV13450	Amino acid sequenc	209	5	55.6	664	21	AAV68925	Amino acids 23-666
137	5	55.6	429	20	AAW73631	Human secreted pro	210	5	55.6	669	20	AAV31939	Human rad17 cell c
138	5	55.6	434	20	AAW73959	Human TP33 protein	211	5	55.6	670	17	AAV79629	Testis-associated
139	5	55.6	442	20	AAV37457	Chlamydia trachoma	212	5	55.6	670	20	AAV31937	Human rad17 cell c
140	5	55.6	453	19	AAW53867	Gravlin polypeptide	213	5	55.6	671	17	AAW85290	Streptococcus faec
141	5	55.6	453	21	AAAB18407	Human gravlin PKA r	214	5	55.6	677	21	AAV68948	Amino acids 29-685
142	5	55.6	461	20	AAV38793	Neisseria gonorrho	215	5	55.6	678	21	AAV68947	Amino acids 28-685
143	5	55.6	461	20	AAV38794	N. gonorrhoeae ant	216	5	55.6	679	21	AAV68946	Amino acids 27-685
144	5	55.6	465	21	AAV81676	Streptococcus pneu	217	5	55.6	680	21	AAV68945	Amino acids 26-685
145	5	55.6	468	18	AAV19722	Mouse p14 (muscle	218	5	55.6	681	21	AAV68944	Amino acids 25-685
146	5	55.6	469	17	AAV24476	Corrected mouse p1	219	5	55.6	682	21	AAV68943	Amino acids 24-685
147	5	55.6	469	17	AAV24476	Corrected mouse p1	220	5	55.6	685	19	AAW80813	Nucleotide sequenc
148	5	55.6	473	20	AAW89835	Protein encoded by	221	5	55.6	685	21	AAW94507	Human delta-2 prot
149	5	55.6	478	10	AAW90459	Alpha-factor profi	222	5	55.6	685	21	AAV68912	Amino acid sequenc
150	5	55.6	478	10	AAW90461	Alpha-factor profi	223	5	55.6	685	21	AAV76912	Human notch ligand
151	5	55.6	480	11	AAW04083	Platlet aggregati	224	5	55.6	686	21	AAV68911	Amino acid sequenc
152	5	55.6	480	14	AAW2867	Platlet aggregati	225	5	55.6	686	21	AAV76911	Mouse notch ligand
153	5	55.6	483	21	AAV79413	Southern copenhea	226	5	55.6	717	10	AAW91933	BI antigen. Toxop
154	5	55.6	486	19	AAW37056	HIV-1 breakthrough	227	5	55.6	730	9	AAW80618	Human Bone Morphog
155	5	55.6	491	15	AAW37057	HIV-1 breakthrough	228	5	55.6	730	18	AAW31669	C-proteinase encod
156	5	55.6	498	15	AAW4154	Human neuronal nic	229	5	55.6	739	18	AAW25790	Gene 036 product of
157	5	55.6	498	18	AAW09027	Neuronal nicotinic	230	5	55.6	764	21	AAV68949	Fusion protein of

231	5	55.6	788	19	AAW75919	C-proteinase seque	304	4	44.4	8	17	AAW95514	Human prostate car
232	5	55.6	797	21	AAW70963	Human Ras signalli	305	4	44.4	8	20	AAW41615	Mammalian ion chan
233	5	55.6	802	21	AAW81746	M. tuberculosis fu	306	4	44.4	8	21	AAW77393	Human I-group O env
234	5	55.6	802	19	AAW64379	Mycobacterium anti	307	4	44.4	8	22	AAW31509	Amino acid sequenc
235	5	55.6	802	20	AAW32063	Mycobacterium tube	308	4	44.4	9	14	AAW44262	Residues 68-76 of
236	5	55.6	802	20	AAW39224	M. tuberculosis fu	309	4	44.4	9	14	AAW42578	HIV envelope neutr
237	5	55.6	802	20	AAW39176	M. tuberculosis fu	310	4	44.4	9	19	AAW79187	M88 murine monoclo
238	5	55.6	802	20	AAW39081	M. tuberculosis fu	311	4	44.4	9	20	AAW94619	Src homology 3 dom
239	5	55.6	802	20	AAW39033	M. tuberculosis fu	312	4	44.4	9	21	AAW98674	W11 derived immuno
240	5	55.6	835	17	AAW96206	Invasin protein.	313	4	44.4	9	21	AAW98680	W11 derived immuno
241	5	55.6	843	13	AAW27744	Extracellular prot	314	4	44.4	9	21	AAW98722	W11 derived immuno
242	5	55.6	870	21	AAW10948	L. mexicana casein	315	4	44.4	9	21	AAW98727	W11 derived immuno
243	5	55.6	933	20	AAW21621	Ligand binding dom	316	4	44.4	9	21	AAW98780	W11 derived immuno
244	5	55.6	933	21	AAW97292	Human progesterone	317	4	44.4	9	22	AAW31508	W11 derived immuno
245	5	55.6	939	21	AAW53192	Mecaca mulatta rha	318	4	44.4	9	22	AAW76154	Amino acid sequenc
246	5	55.6	939	21	AAW67882	Ste4p/Gdelta. Intera	319	4	44.4	10	17	AAW49590	Tumour associated
247	5	55.6	941	21	AAW24422	Human PRO1154 prot	320	4	44.4	10	17	AAW49591	Human leucocyte an
248	5	55.6	941	21	AAW66736	Membrane-bound pro	321	4	44.4	10	17	AAW49592	Human leucocyte an
249	5	55.6	941	22	AAW65259	Human PRO1154 (UNQ	322	4	44.4	10	17	AAW96505	Hepatitis C virus
250	5	55.6	941	22	AAW65007	Human secreted pro	323	4	44.4	10	17	AAW97522	Antigenic peptide,
251	5	55.6	954	18	AAW19752	Yeast inhibitor of	324	4	44.4	10	19	AAW76876	Fusion immunoglobu
252	5	55.6	986	19	AAW13670	C-proteinase encod	325	4	44.4	10	19	AAW76878	Fusion immunoglobu
253	5	55.6	1013	19	AAW61539	Human cardiac/brai	326	4	44.4	10	19	AAW63551	Beta (1 -> 4)-N-ac
254	5	55.6	1013	19	AAW40224	Human tollold-1like	327	4	44.4	10	22	AAW31507	Amino acid sequenc
255	5	55.6	1030	19	AAW53572	Human myosin 11ght	328	4	44.4	10	22	AAW76201	Tumour associated
256	5	55.6	1030	19	AAW41378	Human protein p164	329	4	44.4	10	22	AAW76202	Tumour associated
257	5	55.6	1085	12	AAW11604	PA50.17-alpha/P450	330	4	44.4	11	10	AAW93053	HIV env protein an
258	5	55.6	1148	14	AAW43671	M. leprae rpoB gene	331	4	44.4	11	10	AAW44263	Residues 68-78 of
259	5	55.6	1189	15	AAW56496	TATA-binding prote	332	4	44.4	11	14	AAW54624	Listeria innocua p
260	5	55.6	1213	18	AAW06086	Drosophila TATA-bl	333	4	44.4	11	14	AAW54627	Listeria innocua p
261	5	55.6	1213	18	AAW25029	TATA-binding prote	334	4	44.4	11	17	AAW41809	M13TMD1 mutated fr
262	5	55.6	1252	16	AAW80530	B. sphaericus SLP.	335	4	44.4	11	17	AAW97523	Antigenic peptide,
263	5	55.6	1253	13	AAW28337	SVF4 structural po	336	4	44.4	11	19	AAW50203	Peptide from a kri
264	5	55.6	1268	18	AAW58774	Human breast cance	337	4	44.4	11	21	AAW22972	White shrimp p62 p
265	5	55.6	1268	21	AAW58774	Human insulin rece	338	4	44.4	11	21	AAW09549	Peptide encoded by
266	5	55.6	1297	21	AAW36840	Human insulin rece	339	4	44.4	11	21	AAW02815	BMP mutant chimeri
267	5	55.6	1300	21	AAW36838	Rat insulin recept	340	4	44.4	11	21	AAW92589	Peptide encoded by
268	5	55.6	1315	21	AAW84046	Splice variant of	341	4	44.4	11	22	AAW31506	Amino acid sequenc
269	5	55.6	1353	21	AAW84045	Splice variant of	342	4	44.4	12	13	AAW23605	Peptide able to in
270	5	55.6	1405	21	AAW84044	Amino acid sequenc	343	4	44.4	12	16	AAW83319	KD-binding random
271	5	55.6	1449	21	AAW32825	Murine DEC 205 rec	344	4	44.4	12	22	AAW31505	Amino acid sequenc
272	5	55.6	1484	21	AAW89721	Canine ribosome re	345	4	44.4	13	17	AAW96827	Human snRNP-associ
273	5	55.6	1566	16	AAW79643	Immunodominant ant	346	4	44.4	13	19	AAW70128	Peptide produced b
274	5	55.6	1627	12	AAW12789	M. pneumoniae p1 c	347	4	44.4	13	19	AAW50204	Peptide from a kri
275	5	55.6	1627	15	AAW47911	Mycoplasma pneumon	348	4	44.4	13	21	AAW22973	White shrimp p62 p
276	5	55.6	1627	16	AAW67538	Cytadhesin p1. My	349	4	44.4	13	22	AAW31504	Amino acid sequenc
277	5	55.6	1723	17	AAW00645	Human gravin prote	350	4	44.4	14	13	AAW23604	Peptide able to in
278	5	55.6	1780	19	AAW53863	Human gravin prote	351	4	44.4	14	16	AAW77855	Antigen to raise a
279	5	55.6	1780	21	AAW15380	Amino acid sequenc	352	4	44.4	14	17	AAW92864	Cell adhesion modu
280	5	55.6	1794	21	AAW84040	Extracellular fact	353	4	44.4	14	19	AAW79191	Synthetic peptide
281	5	55.6	1822	13	AAW27745	Novel protein kina	354	4	44.4	14	19	AAW56527	Antigenic HIV-1 pe
282	5	55.6	1920	22	AAW65656	Arabidopsis thalia	355	4	44.4	14	21	AAW22974	White shrimp p62 p
283	5	55.6	2001	22	AAW20062	Human breast cance	356	4	44.4	14	21	AAW98509	Human W11 peptide
284	5	55.6	2343	21	AAW12453	Human HNRCP protei	357	4	44.4	14	21	AAW98510	Human W11 peptide
285	5	55.6	2476	20	AAW67738	Pig p105 zona pell	358	4	44.4	14	22	AAW31503	Amino acid sequenc
286	5	55.6	2476	20	AAW67738	Human down-regulat	359	4	44.4	15	10	AAW90224	Antigenic peptide
287	5	55.6	2785	21	AAW57148	Human ORFX ORF2255	360	4	44.4	15	11	AAW06576	Bovine serum album
288	5	55.6	3266	21	AAW42491	Human breast cance	361	4	44.4	15	16	AAW79629	Endocarditis speci
289	5	55.6	3418	18	AAW19211	Human breast and o	362	4	44.4	15	16	AAW79626	Serum albumin whos
290	5	55.6	3418	18	AAW23287	Human BRCA2 (om12)	363	4	44.4	15	17	AAW15317	Peptide resembling
291	5	55.6	3418	20	AAW04355	Human BRCA2 (om11)	364	4	44.4	15	18	AAW39011	SH2 domain binding
292	5	55.6	3418	20	AAW04356	Human BRCA2 (om13)	365	4	44.4	15	19	AAW78551	Mutated human G-al
293	5	55.6	3418	20	AAW04357	Human BRCA2 (om14)	366	4	44.4	15	20	AAW02215	HIV USA gp120 clea
294	5	55.6	3418	20	AAW04358	Human BRCA2 (om15)	367	4	44.4	15	21	AAW14221	Mutant HIV cleavag
295	5	55.6	3418	20	AAW04354	Human BRCA2 (om11)	368	4	44.4	15	22	AAW31502	Amino acid sequenc
296	5	55.6	3418	20	AAW04354	BRCA2 protein sequ	369	4	44.4	16	18	AAW25388	PlC gamma SH3 doma
297	4	44.4	3418	21	AAW77819	Peptide encoded by	370	4	44.4	16	18	AAW25388	PlC gamma SH3 doma
298	4	44.4	3418	21	AAW09551	BMP mutant chimeri	371	4	44.4	16	18	AAW15570	PNK358. Feline l
299	4	44.4	3418	21	AAW92591	Peptide encoded by	372	4	44.4	16	19	AAW76907	Fusion immunoglobu
300	4	44.4	3418	21	AAW92591	Amino acid sequenc	373	4	44.4	16	19	AAW50206	Peptide from a kri
301	4	44.4	3418	21	AAW92591	Amino acid sequenc	374	4	44.4	16	19	AAW50205	Peptide from a kri
302	4	44.4	3418	21	AAW92591	Sm B/B' epitope 16	375	4	44.4	16	21	AAW22975	White shrimp p62 p
303	4	44.4	3418	21	AAW92591	Antigenic peptide,	376	4	44.4	16	22	AAW31501	Amino acid sequenc

377	4	44.4	16	22	AAB6802	Escherichia coli K
378	4	44.4	17	22	AAB31500	Amino acid sequenc
379	4	44.4	18	19	AAW79190	Synthetic peptide
380	4	44.4	18	20	AAW94621	Src homology 3 dom
381	4	44.4	18	20	AAW84076	S. cerevisiae amin
382	4	44.4	18	22	AAB31499	Amino acid sequenc
383	4	44.4	19	16	AAW67224	gp120 epitope from
384	4	44.4	19	17	AAW02223	CD4:gamma chain fu
385	4	44.4	19	17	AAW02223	CD4:zeta chain fus
386	4	44.4	19	18	AAW33957	Vasopressin type 2
387	4	44.4	19	19	AAW35391	Bifidobacterium 10
388	4	44.4	19	22	AAB31498	Amino acid sequenc
389	4	44.4	20	12	AAR15532	Immunopeptide detl
390	4	44.4	20	14	AAW37689	Synthetic epitope.
391	4	44.4	20	16	AAW68794	Cytotoxic T lympho
392	4	44.4	20	17	AAW08052	HIV peptide #37.
393	4	44.4	20	19	AAW76943	Fusion immunoglob
394	4	44.4	20	19	AAW76875	Fusion immunoglob
395	4	44.4	20	21	AAB03854	Peptide T0120 used
396	4	44.4	20	21	AAW93365	Amino acid sequenc
397	4	44.4	20	21	AAW67262	Protein tyrosine p
398	4	44.4	20	21	AAW66445	HLA-A2-binding HIV
399	4	44.4	20	22	AAB31497	Amino acid sequenc
400	4	44.4	20	22	AAB49381	Simian immunodef
401	4	44.4	21	13	AAW3606	Peptide able to in
402	4	44.4	21	14	AAW31614	Peptide ligand for
403	4	44.4	21	14	AAW37992	Heavy chain hinge
404	4	44.4	21	19	AAW40272	K. oxytoca R-speci
405	4	44.4	21	22	AAB89251	HIV gp120 protein
406	4	44.4	21	22	AAB89252	HIV gp120 protein
407	4	44.4	21	22	AAB89253	HIV gp120 protein
408	4	44.4	21	22	AAB89254	HIV gp120 protein
409	4	44.4	21	22	AAB31496	Amino acid sequenc
410	4	44.4	22	14	AAW30539	HIV discriminatory
411	4	44.4	22	14	AAW31965	BCH-132. Syntheti
412	4	44.4	22	14	AAW41293	Peptide fragment F
413	4	44.4	22	19	AAW67032	Peptide with N-ter
414	4	44.4	22	22	AAB31495	Amino acid sequenc
415	4	44.4	23	14	AAW41294	Peptide fragment F
416	4	44.4	23	14	AAW41332	HIV gp120 epitope.
417	4	44.4	23	15	AAW63849	HIV-1 gp120-40 epi
418	4	44.4	23	16	AAW74046	Rice carbonic-anhy
419	4	44.4	23	16	AAW50207	Peptide from a kit
420	4	44.4	23	19	AAW48108	HIV gp120 residues
421	4	44.4	23	22	AAB31494	Amino acid sequenc
422	4	44.4	24	8	AAW71151	Anti-HIV peptide 6
423	4	44.4	24	9	AAW60750	Cluster of peptid
424	4	44.4	24	14	AAW33838	Peptide fragment PC
425	4	44.4	24	16	AAW66432	PCUS 4-18 (476-49
426	4	44.4	24	19	AAW54933	HIV gp120 envelope
427	4	44.4	24	20	AAW08855	Expression constru
428	4	44.4	24	20	AAW08768	Expression constru
429	4	44.4	24	20	AAW05354	HIV-1 CUVAC pepi
430	4	44.4	24	21	AAW22976	White shrimp p62 p
431	4	44.4	24	22	AAB31493	Amino acid sequenc
432	4	44.4	25	7	AAW60708	HIV virus p18 prot
433	4	44.4	25	13	AAW27336	Peptide corresep. t
434	4	44.4	25	20	AAW81958	Human EMDO-II vari
435	4	44.4	25	22	AAB31492	Hepatitis A virus
436	4	44.4	26	12	AAW1413	HIV discriminatory
437	4	44.4	26	14	AAW30538	HIV discriminatory
438	4	44.4	26	14	AAW30755	Peptide fragment F
439	4	44.4	26	14	AAW41301	HIV gp120 epitope.
440	4	44.4	26	14	AAW44111	Random bioinlylati
441	4	44.4	26	16	AAW5044	Fusion immunoglob
442	4	44.4	26	19	AAW76909	Peptide from a kit
443	4	44.4	26	19	AAW50208	Bioinlylation pep
444	4	44.4	26	19	AAW46647	White shrimp p62 p
445	4	44.4	26	20	AAW29315	Amino acid sequenc
446	4	44.4	26	21	AAW22977	Non-immunogenic am
447	4	44.4	26	22	AAW31491	HIV discriminatory
448	4	44.4	27	10	AAW90819	
449	4	44.4	27	14	AAW30757	
450	4	44.4	27	19	AAW79184	HIV gp120 C5 regio
451	4	44.4	27	20	AAW38474	Human secreted pro
452	4	44.4	27	22	AAB31490	Amino acid sequenc
453	4	44.4	28	15	AAW49685	Sequence of tryptil
454	4	44.4	28	19	AAW21018	Human glial fibrill
455	4	44.4	28	21	AAB38463	Fragment of human
456	4	44.4	28	22	AAB31489	Amino acid sequenc
457	4	44.4	29	21	AAB33288	Pinus radiata tran
458	4	44.4	29	21	AAW68639	Amino acid sequenc
459	4	44.4	29	22	AAB31488	Amino acid sequenc
460	4	44.4	29	22	AAW49392	Simian immunodef
461	4	44.4	29	22	AAW49393	Simian immunodef
462	4	44.4	29	22	AAW49394	Simian immunodef
463	4	44.4	30	12	AAW15700	Human tumour-asso
464	4	44.4	30	19	AAW76911	Fusion immunoglob
465	4	44.4	30	22	AAB31487	Amino acid sequenc
466	4	44.4	31	17	AAW31189	gp120 HTLV-III pep
467	4	44.4	31	18	AAW24744	Finger 2 domain of
468	4	44.4	31	18	AAW19977	Fibronectin fragme
469	4	44.4	31	19	AAW54936	HIV gp120 envelope
470	4	44.4	31	20	AAW48370	Human prostate can
471	4	44.4	31	20	AAW82455	Tomato Cu/Zn SOD e
472	4	44.4	31	20	AAW82988	Human fibronectin
473	4	44.4	31	21	AAW77377	HIV-1 group O env
474	4	44.4	31	21	AAW77378	HIV-1 group O env
475	4	44.4	31	22	AAB31486	Amino acid sequenc
476	4	44.4	32	10	AAW91490	C-terminal portion
477	4	44.4	32	21	AAW98799	WT1 derived immuno
478	4	44.4	32	22	AAB31485	Amino acid sequenc
479	4	44.4	33	16	AAW69686	Hepatitis C virus
480	4	44.4	33	17	AAW89558	Hepatitis C virus
481	4	44.4	33	22	AAB31484	Amino acid sequenc
482	4	44.4	34	14	AAW43826	Anti-GAT antibody
483	4	44.4	34	17	AAW89860	Cytochrome P450 2C
484	4	44.4	34	19	AAW22870	SEQ ID NO. 66 from
485	4	44.4	34	19	AAW48443	Human p53 proline-
486	4	44.4	34	22	AAB31481	Amino acid sequenc
487	4	44.4	35	16	AAW67556	SIV-MND gp120 V3 I
488	4	44.4	35	19	AAW670114	Peptide produced b
489	4	44.4	35	20	AAW11421	Human 5' EST sece
490	4	44.4	35	21	AAB38462	Fragment of human
491	4	44.4	35	21	AAW67806	Peptide #206 for d
492	4	44.4	35	21	AAW67808	Peptide #208 for d
493	4	44.4	35	21	AAW67809	Peptide #209 for d
494	4	44.4	36	21	AAW38414	Fragment of human
495	4	44.4	37	18	AAW40369	Human breast cance
496	4	44.4	37	18	AAW30366	Fragment #2 of bon
497	4	44.4	37	20	AAW25880	Human secreted pro
498	4	44.4	37	21	AAW44604	Human secreted pro
499	4	44.4	37	21	AAW09516	Human Op-1 finger
500	4	44.4	37	21	AAW09559	Op-1 mutant protei
501	4	44.4	37	21	AAW09562	Human Op-1 mutant
502	4	44.4	37	21	AAW09563	Human Op-1 mutant
503	4	44.4	37	21	AAW09565	Human Op-1 mutant
504	4	44.4	37	21	AAW09566	Human Op-1 mutant
505	4	44.4	37	21	AAW09573	Human Op-1 mutant
506	4	44.4	37	21	AAW84032	Amino acid sequenc
507	4	44.4	38	14	AAW41075	HIV-1 gp120 C-term
508	4	44.4	38	15	AAW58457	TSAR binding domai
509	4	44.4	38	17	AAW95453	Antigen-binding pe
510	4	44.4	38	20	AAW42830	Erythropoietin rec
511	4	44.4	38	20	AAW42835	Non-EPOR-binding c
512	4	44.4	38	20	AAW74072	Human HPT-1 bindin
513	4	44.4	38	21	AAW55741	Arabidopsis thaila
514	4	44.4	39	16	AAW66436	PCUS 4-18 (476-49
515	4	44.4	39	16	AAW66410	PCUS 4-18 (476-49
516	4	44.4	39	18	AAW19025	HIV envelope glyco
517	4	44.4	39	19	AAW75198	Human secreted pro
518	4	44.4	39	20	AAW05347	HIV-1 CUVAC pepi
519	4	44.4	39	20	AAW05340	HIV-1 CUVAC pepi
520	4	44.4	39	22	AAW59869	HIV-1 C5 region of
521	4	44.4	41	19	AAW22869	SEQ ID NO. 65 from
522	4	44.4	41	20	AAW12310	Human 5' EST sece



523	4	44.4	42	18	AAW19055	SP410-BAL, Immunoge
524	4	44.4	42	18	AAW06501	Cytomegalovirus la
525	4	44.4	42	20	AAV12163	Human 5' EST seque
526	4	44.4	43	20	AAV12848	Human 5' EST seque
527	4	44.4	43	10	AAW0818	Non-Immunogenic am
528	4	44.4	45	22	AAW87380	Human gene 39 enco
529	4	44.4	46	19	AAW79183	Fusion immunoglobu
530	4	44.4	46	21	AAW28505	Zea mays protein f
531	4	44.4	47	20	AAV36545	Fragment of human
532	4	44.4	47	21	AAW45201	Arabidopsis thalia
533	4	44.4	49	19	AAW46754	Exon trap L75917.
534	4	44.4	50	16	AAW1628	HIV gp120-C5. Syn
535	4	44.4	50	17	AAW98476	Cytochrome P450 2C
536	4	44.4	50	21	AAW09092	Hepatitis GB virus
537	4	44.4	51	16	AAW77798	Transactivation do
538	4	44.4	51	16	AAW77799	Transactivation do
539	4	44.4	51	20	AAV48336	Human prostate can
540	4	44.4	51	20	AAV36715	Fragment of human
541	4	44.4	51	21	AAW16822	Bacteriophage Dp-1
542	4	44.4	51	21	AAW09225	Hepatitis GB virus
543	4	44.4	52	10	AAW91361	Amino acids of the
544	4	44.4	52	21	AAW22970	White shrimp multi
545	4	44.4	52	21	AAW22971	White shrimp multi
546	4	44.4	53	17	AAW88330	Cytochrome P450 2C
547	4	44.4	53	19	AAW22898	SEQ ID NO. 94 from
548	4	44.4	53	20	AAV30706	Amino acid sequenc
549	4	44.4	54	17	AAW05836	Mouse M-Serrate-1
550	4	44.4	54	19	AAW22867	SEQ ID NO. 63 from
551	4	44.4	54	19	AAW22868	SEQ ID NO. 64 from
552	4	44.4	54	21	AAW59095	Breast and ovarian
553	4	44.4	54	21	AAW00997	Human secreted pro
554	4	44.4	54	21	AAV59595	Mouse Serrate prot
555	4	44.4	55	8	AAW70360	Sequence of immuno
556	4	44.4	55	19	AAW22866	SEQ ID NO. 62 from
557	4	44.4	55	19	AAW22865	SEQ ID NO. 61 from
558	4	44.4	55	20	AAW26920	Fely envelope glyco
559	4	44.4	56	20	AAW02478	Clone selected aft
560	4	44.4	56	21	AAW54442	Human pancreatic c
561	4	44.4	56	22	AAW65854	Human TANCO 281 p
562	4	44.4	56	22	AAW65862	Murine TANCO 281 p
563	4	44.4	57	21	AAW54473	Zea mays protein f
564	4	44.4	58	10	AAW91360	Amino acids 455-51
565	4	44.4	59	16	AAW73922	B.catarrahalis CD e
566	4	44.4	59	17	AAW04957	Partial F. anguioi
567	4	44.4	59	21	AAW58213	Arabidopsis thalia
568	4	44.4	59	21	AAW59293	Arabidopsis thalia
569	4	44.4	59	21	AAW00608	Human secreted pro
570	4	44.4	59	22	AAW96790	Maize sucrose non-
571	4	44.4	59	22	AAW90628	Human secreted pro
572	4	44.4	60	18	AAW15568	PRN60. Feline leu
573	4	44.4	60	19	AAW20977	Human glial fibrin
574	4	44.4	60	19	AAW2727	Proline rich pepti
575	4	44.4	60	20	AAV11876	Human 5' EST seque
576	4	44.4	60	21	AAW6896	Arabidopsis thalia
577	4	44.4	61	14	AAW39350	Basic amphipathic
578	4	44.4	61	21	AAW09627	Arabidopsis thalia
579	4	44.4	61	21	AAW02506	Human secreted pro
580	4	44.4	62	20	AAV19648	SEQ ID NO 366 from
581	4	44.4	63	21	AAW51667	Human secreted pro
582	4	44.4	63	21	AAW59038	Arabidopsis thalia
583	4	44.4	64	18	AAW11269	Type III AFP varia
584	4	44.4	64	20	AAW38898	Neisseria meningit
585	4	44.4	64	21	AAW45200	Arabidopsis thalia
586	4	44.4	64	21	AAW55463	Arabidopsis thalia
587	4	44.4	64	21	AAV90583	Conus striatus thal
588	4	44.4	65	21	AAW38595	Human secreted pro
589	4	44.4	65	22	AAW76042	Human skin cell pr
590	4	44.4	65	22	AAW79213	Corynebacterium gl
591	4	44.4	65	22	AAW55981	Skin cell protein,
592	4	44.4	67	19	AAW83944	Human secreted pro
593	4	44.4	67	19	AAW57058	Human prostate can
594	4	44.4	68	21	AAW65384	Human 5' EST relat
595	4	44.4	69	18	AAW32434	Mycobacterium tube
596	4	44.4	69	18	AAW32366	Mycobacterium tube
597	4	44.4	69	18	AAW12764	A-lineage conotoxi
598	4	44.4	69	19	AAW81669	M. tuberculosis im
599	4	44.4	69	19	AAW64306	Mycobacterium tube
600	4	44.4	69	20	AAV39108	M. tuberculosis re
601	4	44.4	69	20	AAV38971	M. tuberculosis re
602	4	44.4	69	21	AAW55153	Arabidopsis thalia
603	4	44.4	70	20	AAV28247	Human secreted pro
604	4	44.4	70	20	AAW93960	Human 53BP2:IP-2 p
605	4	44.4	70	21	AAW53453	Human colon cancer
606	4	44.4	70	21	AAW03849	Protein fragment #
607	4	44.4	71	19	AAV14462	Human secreted pro
608	4	44.4	72	20	AAV22887	SEQ ID NO. 83 from
609	4	44.4	73	22	AAW70028	Mammalian CD4 IL-1
610	4	44.4	73	22	AAW70018	Human CD4 D4 domai
611	4	44.4	73	22	AAW70025	Human CD4 D4 domai
612	4	44.4	73	22	AAW70026	Human CD4 D4 domai
613	4	44.4	73	22	AAW70027	Human CD4 D4 domai
614	4	44.4	73	22	AAW70029	Human CD4 D4 domai
615	4	44.4	74	20	AAV14475	Human CD4 D4 domai
616	4	44.4	74	20	AAV03781	Fragment of human
617	4	44.4	74	20	AAW56214	S. aureus polypept
618	4	44.4	74	21	AAW37064	Human secreted pro
619	4	44.4	74	21	AAW5678	Arabidopsis thalia
620	4	44.4	74	21	AAW60223	Arabidopsis thalia
621	4	44.4	74	21	AAV50946	Arabidopsis thalia
622	4	44.4	75	20	AAV35810	Human adult thymus
623	4	44.4	75	20	AAV12951	Chlamydia pneumoni
624	4	44.4	75	21	AAW18074	Amino acid sequenc
625	4	44.4	75	21	AAW90584	Arabidopsis thalia
626	4	44.4	75	21	AAV90587	Conus striatus thal
627	4	44.4	76	21	AAW55739	Conus sulcatus kap
628	4	44.4	76	21	AAW01671	Arabidopsis thalia
629	4	44.4	76	22	AAW65860	Human secreted pro
630	4	44.4	77	19	AAW66001	Murine TANCO 281 c
631	4	44.4	77	19	AAW25354	S. pneumoniae derl
632	4	44.4	77	21	AAW46895	Pinus radiata cell
633	4	44.4	78	21	AAW52108	Arabidopsis thalia
634	4	44.4	78	21	AAW55152	Human secreted pro
635	4	44.4	78	22	AAW87389	Human gene 48 enco
636	4	44.4	78	22	AAW87412	Human gene 48 enco
637	4	44.4	79	19	AAW22863	SEQ ID NO. 59 from
638	4	44.4	79	19	AAW22864	SEQ ID NO. 60 from
639	4	44.4	79	20	AAV73910	Human prostate tum
640	4	44.4	79	21	AAW54421	Human pancreatic c
641	4	44.4	80	18	AAW27747	Staphylococcus aur
642	4	44.4	80	18	AAW12765	A-lineage conotoxi
643	4	44.4	80	21	AAW60234	Arabidopsis thalia
644	4	44.4	81	20	AAW60234	Human endometriu
645	4	44.4	81	20	AAW28624	Human secreted pro
646	4	44.4	81	20	AAV73395	Human secreted pro
647	4	44.4	82	21	AAW37388	Human secreted pro
648	4	44.4	82	21	AAW18950	Human secreted pro
649	4	44.4	82	21	AAW22927	Peptide derived fr
650	4	44.4	82	21	AAW79629	Calreticulin Golgi
651	4	44.4	83	21	AAW54450	Golgi target seque
652	4	44.4	84	19	AAW42038	Human pancreatic c
653	4	44.4	85	21	AAW02769	Mouse procorrtstat
654	4	44.4	85	21	AAV91606	Human secreted pro
655	4	44.4	86	19	AAW38421	Human secreted pro
656	4	44.4	86	20	AAV73904	Proteolipid protei
657	4	44.4	86	20	AAV73929	Human prostate tum
658	4	44.4	86	20	AAV11815	Human 5' EST seque
659	4	44.4	87	21	AAW44140	Arabidopsis thalia
660	4	44.4	88	21	AAW04407	Arabidopsis thalia
661	4	44.4	88	21	AAW28504	Zea mays protein f
662	4	44.4	88	21	AAW33200	Zea mays protein f
663	4	44.4	88	21	AAW35339	Zea mays protein f
664	4	44.4	88	21	AAW54249	Arabidopsis thalia
665	4	44.4	89	20	AAW12949	Amino acid sequenc
666	4	44.4	89	21	AAW04406	Arabidopsis thalia
667	4	44.4	89	21	AAW28503	Zea mays protein f
668	4	44.4	89	21	AAW54248	Arabidopsis thalia

669	4	44.4	89	21	AA601533	Human secreted pro	742	4	44.4	102	21	AA602800	Human Op-1 amino a
670	4	44.4	89	21	AA601562	Human secreted pro	743	4	44.4	102	21	AA692569	Op-1 finger-1-heal
671	4	44.4	90	21	AA632306	Corn acid triacylg	744	4	44.4	102	22	AA663506	Human secreted pro
672	4	44.4	91	20	AA114146	Human 5' EST seque	745	4	44.4	103	20	AA30835	Human secreted pro
673	4	44.4	91	21	AA40613	Human ORFX ORP377	746	4	44.4	103	21	AA40704	Human ORFX ORP468
674	4	44.4	91	21	AA42378	Human ORFX ORP2142	747	4	44.4	104	9	AA680104	Sequence encoded b
675	4	44.4	91	21	AA627140	Zea mays protein f	748	4	44.4	104	19	AA422878	SEQ ID NO. 74 from
676	4	44.4	92	20	AA60469	Human normal blad	749	4	44.4	104	21	AA641199	Human ORFX ORP63
677	4	44.4	92	21	AA626934	Zea mays protein f	750	4	44.4	104	21	AA642842	Human ORFX ORP2606
678	4	44.4	92	21	AA633234	Zea mays protein f	751	4	44.4	104	21	AA624955	Arabidopsis thalia
679	4	44.4	92	21	AA633943	Arabidopsis thalia	752	4	44.4	104	21	AA694933	Human secreted pro
680	4	44.4	92	21	AA654495	Zea mays protein f	753	4	44.4	105	19	AA122856	SEQ ID NO. 52 from
681	4	44.4	93	19	AA122860	SEQ ID NO. 56 from	754	4	44.4	105	19	AA122857	SEQ ID NO. 53 from
682	4	44.4	93	19	AA675090	Human secreted pro	755	4	44.4	105	19	AA422858	SEQ ID NO. 54 from
683	4	44.4	93	21	AA604405	Arabidopsis thalia	756	4	44.4	105	19	AA422859	SEQ ID NO. 55 from
684	4	44.4	93	21	AA633473	Arabidopsis thalia	757	4	44.4	105	19	AA40579	Human lambda CL do
685	4	44.4	93	21	AA654247	Arabidopsis thalia	758	4	44.4	105	20	AA38473	Human secreted pro
686	4	44.4	94	17	AA689859	Cytochrome P450 2C	759	4	44.4	105	20	AA108746	Human secreted pro
687	4	44.4	94	19	AA122861	SEQ ID NO. 57 from	760	4	44.4	105	20	AA692426	Human lambda-CL do
688	4	44.4	94	19	AA675151	Human secreted pro	761	4	44.4	105	21	AA627001	Human lambda CL do
689	4	44.4	94	19	AA675152	Human secreted pro	762	4	44.4	105	21	AA612168	Arabidopsis thalia
690	4	44.4	94	21	AA651698	Human secreted pro	763	4	44.4	106	20	AA670803	Amino acid sequenc
691	4	44.4	94	21	AA608039	Arabidopsis thalia	764	4	44.4	106	21	AA640717	Human ORFX ORP481
692	4	44.4	94	21	AA626159	Zea mays protein f	765	4	44.4	106	21	AA637647	Arabidopsis thalia
693	4	44.4	94	21	AA637331	Arabidopsis thalia	766	4	44.4	106	21	AA601566	Human secreted pro
694	4	44.4	94	22	AA664488	Human secreted pro	767	4	44.4	106	21	AA692193	Human Iggl lambda
695	4	44.4	95	18	AA620297	H. pylori cytoplas	768	4	44.4	107	19	AA675784	Human lymphocyte s
696	4	44.4	95	20	AA674161	Human prostate tum	769	4	44.4	107	21	AA642526	Human ORFX ORP2290
697	4	44.4	95	21	AA641126	Zea mays protein f	770	4	44.4	107	21	AA609358	Arabidopsis thalia
698	4	44.4	95	21	AA657203	Arabidopsis thalia	771	4	44.4	108	20	AA129241	Amino acid sequenc
699	4	44.4	95	21	AA659156	Arabidopsis thalia	772	4	44.4	108	20	AA105478	C. albicans Rbt1 p
700	4	44.4	95	21	AA191621	Human secreted pro	773	4	44.4	108	21	AA618536	Zea mays protein f
701	4	44.4	96	21	AA632768	Eucalyptus grandis	774	4	44.4	108	21	AA603520	Human secreted pro
702	4	44.4	96	21	AA606013	Arabidopsis thalia	775	4	44.4	109	14	AA639729	First type III fib
703	4	44.4	96	21	AA608038	Arabidopsis thalia	776	4	44.4	109	18	AA637348	Immunoglobulin C-1
704	4	44.4	96	21	AA637330	Arabidopsis thalia	777	4	44.4	109	19	AA657189	Fibronectin first
705	4	44.4	97	14	AA653359	Osteogenic protein	778	4	44.4	109	19	AA642034	Mouse preprocortis
706	4	44.4	97	14	AA653387	Murine osteogenic	779	4	44.4	109	20	AA428574	Secreted peptide c
707	4	44.4	97	19	AA638643	Streptococcus pneu	780	4	44.4	109	20	AA345593	Chlamydia pneumoni
708	4	44.4	97	20	AA627248	C. elegans CED-6 p	781	4	44.4	109	20	AA688817	Polypeptide fragme
709	4	44.4	97	20	AA689659	Human osteogenic p	782	4	44.4	109	20	AA682982	Human fibronectin
710	4	44.4	97	20	AA695444	Conserved 6 cystei	783	4	44.4	110	12	AA612387	Amplified human ap
711	4	44.4	97	21	AA633109	Pinus radiata trun	784	4	44.4	110	21	AA624425	Human PRO1274 prot
712	4	44.4	97	21	AA633072	Arabidopsis thalia	785	4	44.4	110	21	AA615654	Arabidopsis thalia
713	4	44.4	97	21	AA648926	Arabidopsis thalia	786	4	44.4	110	21	AA618889	Zea mays protein f
714	4	44.4	97	21	AA690944	Comomoras testoste	787	4	44.4	110	21	AA699380	Human PRO1274 (UNQ
715	4	44.4	97	22	AA665852	Human OSTO281 cy	788	4	44.4	110	21	AA687273	Human signal pepti
716	4	44.4	98	10	AA695681	Human osteogenic p	789	4	44.4	110	22	AA666129	Protein of the Inv
717	4	44.4	99	21	AA641554	Arabidopsis thalia	790	4	44.4	111	16	AA666317	Human Immunoglobul
718	4	44.4	100	15	AA647238	Wild-type feline H	791	4	44.4	111	18	AA628276	Amino acid sequenc
719	4	44.4	100	21	AA627328	Arabidopsis thalia	792	4	44.4	111	21	AA612314	Zea mays protein f
720	4	44.4	100	21	AA636857	Human secreted thalia	793	4	44.4	111	21	AA603420	Human secreted pro
721	4	44.4	100	21	AA600536	Human secreted pro	794	4	44.4	111	21	AA692027	Human bone morphog
722	4	44.4	100	22	AA681811	Streptococcus pneu	795	4	44.4	112	21	AA619019	Zea mays protein f
723	4	44.4	100	22	AA631970	Rice glutamate 1-s	796	4	44.4	112	21	AA657235	Arabidopsis thalia
724	4	44.4	101	20	AA135775	Chlamydia pneumoni	797	4	44.4	113	17	AA696560	Hepatitis C virus
725	4	44.4	101	20	AA612050	Human 5' EST seque	798	4	44.4	113	17	AA696561	Hepatitis C virus
726	4	44.4	102	10	AA695682	Human osteogenic p	799	4	44.4	113	17	AA696562	Hepatitis C virus
727	4	44.4	102	14	AA653360	Osteogenic protein	800	4	44.4	113	17	AA696564	Hepatitis C virus
728	4	44.4	102	18	AA636889	Human osteogenic p	801	4	44.4	113	21	AA640765	Human ORFX ORP529
729	4	44.4	102	18	AA636897	Mutant human osteo	802	4	44.4	113	21	AA615925	Arabidopsis thalia
730	4	44.4	102	18	AA636872	Human osteogenic p	803	4	44.4	113	21	AA660083	Arabidopsis thalia
731	4	44.4	102	18	AA636873	Mouse osteogenic p	804	4	44.4	113	21	AA660682	Arabidopsis thalia
732	4	44.4	102	19	AA678901	Human UNC-5 homolo	805	4	44.4	113	21	AA600281	Human secreted pro
733	4	44.4	102	20	AA636828	Chlamydia trachoma	806	4	44.4	114	14	AA65361	N-terminally trunc
734	4	44.4	102	20	AA16708	WO9914235 Seq ID N	807	4	44.4	114	20	AA634698	Chlamydia pneumoni
735	4	44.4	102	20	AA119744	SEQ ID NO 462 from	808	4	44.4	114	20	AA695454	N-terminally trunc
736	4	44.4	102	20	AA689697	Human osteogenic p	809	4	44.4	114	21	AA608037	Arabidopsis thalia
737	4	44.4	102	20	AA695443	Conserved 7 cystei	810	4	44.4	114	21	AA637329	Arabidopsis thalia
738	4	44.4	102	21	AA654387	Human pancreatic c	811	4	44.4	114	21	AA675983	Murine skin cell p
739	4	44.4	102	21	AA606012	Arabidopsis thalia	812	4	44.4	114	21	AA676033	Murine skin cell p
740	4	44.4	102	21	AA644007	Zea mays protein f	813	4	44.4	114	22	AA655822	Skin cell protein,
741	4	44.4	102	21	AA609534	Human Op-1 C-termi	814	4	44.4	114	22	AA655972	skin cell protein,

815	4	44.4	115	18	AAW12437	OP-1 based morphon
816	4	44.4	115	20	AAV42461	Rat granulosa prepr
817	4	44.4	115	20	AAV39287	Human phosphodiester
818	4	44.4	115	20	AAV12133	Human 5' EST secre
819	4	44.4	115	20	AAW97653	Wheat SUG1 polypep
820	4	44.4	115	21	AAAG58270	Arabidopsis thalia
821	4	44.4	116	14	AAAR5362	N-terminally trunc
822	4	44.4	116	16	AAAR78728	Rnf gene product d
823	4	44.4	116	20	AAW95453	N-terminally trunc
824	4	44.4	116	21	AAW56080	Human secreted pro
825	4	44.4	117	14	AAAR5363	N-terminally trunc
826	4	44.4	117	18	AAV11162	S. pneumoniae prot
827	4	44.4	117	20	AAV74203	Human prostate tum
828	4	44.4	117	20	AAV66146	Human bladder tumo
829	4	44.4	117	20	AAV37327	Amino acid sequenc
830	4	44.4	117	20	AAV35533	Chlamydia pneumoni
831	4	44.4	117	20	AAW92252	Bu172 polypeptide
832	4	44.4	117	20	AAW95452	N-terminally trunc
833	4	44.4	117	21	AAAG24952	Arabidopsis thalia
834	4	44.4	117	21	AAAB09548	Trypsin truncated
835	4	44.4	117	21	AAAB02814	Human trypsin trun
836	4	44.4	117	21	AAV92595	Trypsin truncated
837	4	44.4	117	21	AAV70526	Maize plastid targ
838	4	44.4	118	16	AAAR7184	Chimeric protein g
839	4	44.4	118	21	AAAB57065	Human prostate can
840	4	44.4	118	21	AAAG57234	Arabidopsis thalia
841	4	44.4	119	14	AAAR5364	N-terminally trunc
842	4	44.4	119	20	AAV34585	Chlamydia pneumoni
843	4	44.4	119	20	AAW95451	N-terminally trunc
844	4	44.4	119	21	AAAG03931	Human secreted pro
845	4	44.4	119	22	AAAB08893	Human secreted pro
846	4	44.4	119	22	AAAB20301	Arabidopsis apopto
847	4	44.4	119	22	AAAB62005	B-myb protein frag
848	4	44.4	120	21	AAAG36031	Zea mays protein f
849	4	44.4	120	21	AAAG36653	Arabidopsis thalia
850	4	44.4	120	22	AAAB47126	CDIF-4, Incyte ID
851	4	44.4	121	19	AAAB69979	Human microtubule
852	4	44.4	121	21	AAAB53900	Human colon cancer
853	4	44.4	121	21	AAAG33951	Arabidopsis thalia
854	4	44.4	122	10	AAAP93697	Protein encoded by
855	4	44.4	122	13	AAAR22381	Antigen tg-3e. EI
856	4	44.4	122	16	AAAR75407	HIV-3 strain AMT70
857	4	44.4	122	20	AAV38835	Neisseria meningit
858	4	44.4	122	21	AAAB42096	Human ORFX ORF1860
859	4	44.4	122	21	AAAG17953	Arabidopsis thalia
860	4	44.4	123	21	AAAB40679	Human ORFX ORF43
861	4	44.4	123	21	AAV50920	Human fetal brain
862	4	44.4	124	15	AAAR54336	Anti-HIV gp120 imm
863	4	44.4	124	17	AAAM01310	VH region of HIV n
864	4	44.4	124	18	AAAM33607	Human secreted pro
865	4	44.4	124	19	AAAM56732	Nucleus specific
866	4	44.4	124	20	AAV74122	Human prostate tum
867	4	44.4	124	20	AAV60074	Human endometrium
868	4	44.4	124	21	AAAB10230	Human fetal kidney
869	4	44.4	124	21	AAAG27327	Arabidopsis thalia
870	4	44.4	124	21	AAAG34082	Zea mays protein f
871	4	44.4	124	21	AAAG41932	Arabidopsis thalia
872	4	44.4	124	21	AAAG58520	Zea mays protein f
873	4	44.4	124	21	AAAG00852	Human secreted pro
874	4	44.4	124	21	AAV95162	Anti-gp120 HCl3 he
875	4	44.4	124	21	AAV98271	Anti-gp120 HCl3 he
876	4	44.4	125	20	AAV04704	Mouse Rpppa amino
877	4	44.4	125	22	AAAB74775	Human MPJ3 protein
878	4	44.4	125	22	AAAB35230	M tuberculosis RV3
879	4	44.4	126	10	AAAP90479	Chimeric monoclonal
880	4	44.4	126	21	AAAG03617	Human secreted pro
881	4	44.4	127	13	AAAR29010	p146-k3 protein pr
882	4	44.4	127	18	AAAM08945	Kappa light chain
883	4	44.4	127	19	AAAM44176	Monoclonal antibod
884	4	44.4	127	21	AAAG28002	Arabidopsis thalia
885	4	44.4	127	21	AAAG41891	Arabidopsis thalia
886	4	44.4	127	22	AAAB8497	Haemophilus influe
887	4	44.4	128	20	AAW78449	H. annuus pollen al
888	4	44.4	128	20	AAW78450	H. annuus pollen al
889	4	44.4	128	20	AAW78451	H. annuus pollen al
890	4	44.4	128	20	AAW78448	H. annuus pollen al
891	4	44.4	128	21	AAAG05701	Arabidopsis thalia
892	4	44.4	128	21	AAAG22977	Arabidopsis thalia
893	4	44.4	128	21	AAAG34081	Zea mays protein f
894	4	44.4	128	21	AAAG35338	Zea mays protein f
895	4	44.4	128	21	AAAG42644	Arabidopsis thalia
896	4	44.4	128	21	AAAG52278	Arabidopsis thalia
897	4	44.4	129	21	AAAB01792	Rice M10 homologue
898	4	44.4	129	21	AAAG36652	Arabidopsis thalia
899	4	44.4	130	9	AAAP83147	Corresponds to p14
900	4	44.4	130	19	AAAM70377	Anti-human CD23 6G
901	4	44.4	130	20	AAV73825	Human prostate tum
902	4	44.4	130	21	AAAG33140	Arabidopsis thalia
903	4	44.4	130	21	AAAG58587	Arabidopsis thalia
904	4	44.4	130	21	AAAG02970	Human secreted pro
905	4	44.4	130	21	AAV81654	Streptococcus pneu
906	4	44.4	130	22	AAAB70790	N. magdali bacter
907	4	44.4	131	21	AAAG10968	Arabidopsis thalia
908	4	44.4	131	21	AAAG36856	Arabidopsis thalia
909	4	44.4	131	22	AAAB64639	Human secreted pro
910	4	44.4	132	14	AAAR53365	N. alata arabinoga
911	4	44.4	132	16	AAAR75510	Human RPPpa amino
912	4	44.4	132	20	AAV04697	Human 5' EST secre
913	4	44.4	132	20	AAV12419	N-terminally trunc
914	4	44.4	132	20	AAW95450	Human ORFX ORF27
915	4	44.4	132	21	AAAB41063	Arabidopsis thalia
916	4	44.4	132	21	AAAG08058	Arabidopsis thalia
917	4	44.4	132	21	AAAG43226	Human secreted pro
918	4	44.4	133	20	AAV60073	Human endometrium
919	4	44.4	133	20	AAV37947	Amino acid sequenc
920	4	44.4	133	20	AAW78447	H. annuus pollen al
921	4	44.4	133	20	AAW78441	H. annuus pollen al
922	4	44.4	133	21	AAAB41454	Human ORFX ORF1218
923	4	44.4	133	21	AAAG04077	Human secreted pro
924	4	44.4	134	21	AAAB41939	Human ORFX ORF1703
925	4	44.4	134	22	AAAB69970	M. jannaschl EF-P
926	4	44.4	134	22	AAAB79820	Corynebacterium gl
927	4	44.4	135	19	AAAB62642	E. coli serine protea
928	4	44.4	135	21	AAAB41247	Human ORFX ORF1011
929	4	44.4	135	21	AAAG10861	Arabidopsis thalia
930	4	44.4	135	21	AAAG26078	Zea mays protein f
931	4	44.4	135	21	AAAG03345	Human secreted pro
932	4	44.4	136	20	AAV37091	Amino acid sequenc
933	4	44.4	136	21	AAAG33199	Zea mays protein f
934	4	44.4	136	21	AAAG33942	Arabidopsis thalia
935	4	44.4	136	21	AAAG36651	Arabidopsis thalia
936	4	44.4	137	20	AAV29171	Amino acid sequenc
937	4	44.4	137	20	AAV37129	Amino acid sequenc
938	4	44.4	137	21	AAAG10860	Arabidopsis thalia
939	4	44.4	137	21	AAAG57233	Arabidopsis thalia
940	4	44.4	138	19	AAAM47361	Human ORFX ORF316
941	4	44.4	138	19	AAAM42019	Arabidopsis thalia
942	4	44.4	138	19	AAV08297	Zea mays protein f
943	4	44.4	138	20	AAV04923	Human growth facto
944	4	44.4	138	20	AAAB43352	Mycoacterium spec
945	4	44.4	138	21	AAAG12821	Human ORFX ORF316
946	4	44.4	138	21	AAAG33233	Arabidopsis thalia
947	4	44.4	138	21	AAAG36403	Zea mays protein f
948	4	44.4	139	13	AAAR23187	Arabidopsis thalia
949	4	44.4	139	13	AAAR23187	Osteogenic protein
950	4	44.4	139	13	AAAR27285	Mature human oste
951	4	44.4	139	13	AAAR27286	Mature murine oste
952	4	44.4	139	14	AAAR53366	N-terminally trunc
953	4	44.4	139	14	AAAR53386	Murine osteogenic
954	4	44.4	139	14	AAAR33398	Human mature OP-1
955	4	44.4	139	14	AAAR33399	Mouse mature OP-1
956	4	44.4	139	14	AAAR33921	Morphogen hOP-1
957	4	44.4	139	14	AAAR33922	Morphogen hOP-1
958	4	44.4	139	15	AAAR31467	Mature p3 OR 31-34
959	4	44.4	139	15	AAAR46724	Human mature osteo
960	4	44.4	139	15	AAAR46733	Mouse osteogenic p
	4	44.4	139	15	AAAR60967	Mature human OP-1

961	4	44.4	139	15	AAR60968	Murine OP-1. Mus
962	4	44.4	139	16	AAR71974	BMP-7. Homo sapie
963	4	44.4	139	17	AAW00222	Mouse mature osteo
964	4	44.4	139	17	AAW00221	Human mature osteo
965	4	44.4	139	18	AAW40180	Human hOP1 mature
966	4	44.4	139	18	AAW40181	Mouse hOP1 mature
967	4	44.4	139	18	AAW36871	Mature protein aml
968	4	44.4	139	19	AAW63003	Mouse OP1 protein
969	4	44.4	139	19	AAW54064	Bone morphogenetic
970	4	44.4	139	20	AAV13210	Human secreted pro
971	4	44.4	139	20	AAW95449	Mature human osteo
972	4	44.4	139	20	AAW84215	Bone morphogenetic
973	4	44.4	139	21	AAW82392	Arabidopsis thalia
974	4	44.4	139	21	AAW36208	Arabidopsis thalia
975	4	44.4	139	21	AAW09547	Mature OP-1 mutant
976	4	44.4	139	21	AAW02813	Human mature H223
977	4	44.4	139	21	AAW92594	Mature mutant OP-1
978	4	44.4	139	21	AAW70755	Mature modified hu
979	4	44.4	139	21	AAW57218	Human osteogenic p
980	4	44.4	139	21	AAW57219	Mouse osteogenic p
981	4	44.4	140	19	AAW98699	H. pylori GHP0 693
982	4	44.4	140	20	AAW06111	Nelson Bay virus p
983	4	44.4	140	21	AAW18399	Arabidopsis thalia
984	4	44.4	140	21	AAW33950	Arabidopsis thalia
985	4	44.4	140	21	AAW39677	Arabidopsis thalia
986	4	44.4	140	21	AAW02053	Human secreted pro
987	4	44.4	140	21	AAW92593	Mutant OP-1 const
988	4	44.4	140	22	AAW78968	C. glutamicum SPT
989	4	44.4	141	21	AAW03951	Human mesenchymal
990	4	44.4	141	21	AAW43355	Human ORFX ORF3119
991	4	44.4	141	21	AAW18094	Arabidopsis thalia
992	4	44.4	141	21	AAW28000	Arabidopsis thalia
993	4	44.4	141	21	AAW41890	Arabidopsis thalia
994	4	44.4	141	21	AAW95429	Human calcium chan
995	4	44.4	142	13	AAW28985	Thyroid N-acetyl-g
996	4	44.4	142	21	AAW41796	Human ORFX ORF1560
997	4	44.4	142	21	AAW12791	Clostridium cellu
998	4	44.4	142	21	AAW08989	Arabidopsis thalia
999	4	44.4	142	21	AAW53843	Arabidopsis thalia
1000	4	44.4	142	22	AAW74776	Rat MP3 protein.

## ALIGNMENTS

```

RESULT 1
ID AAR45168 standard; protein: 9 AA.
XX
XX AAR45168:
XX
XX
XX 16-JUN-1994 (first entry)
XX
XX
XX Listeria p60 peptide epitope.
XX
XX Listeria monocytogenes; antibodies; immunoassay; conjugate.
XX
XX Synthetic.
XX
XX DE4318450-A.
XX
XX 16-DEC-1993.
XX
XX
XX 03-JUN-1993; 93DE-4318450.
XX
XX
XX 11-JUN-1992; 92DE-4219111.
XX
XX 25-NOV-1992; 92DE-4239567.
XX
XX (MERE ) MERCK PATENT GMBH.
XX
XX PA
XX PI Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;
XX PI Hofmann G, Bubert A, Goebel W, Koehler S;
XX

```

```

DR WPI; 1993-406956/51.
XX
XX New primers for PCR detection of Listeria - including individual
PT species, also new peptide(s) for raising antibodies for
PT immunochemical detection
XX
XX
XX Disclosure; Fig 2; 19pp; German.
PS
XX
XX The sequence is that of a Listeria p60 peptide epitope which
CC which may be used in the prodn. of antibodies for the detection
CC of Listeria by immunoassay (partic. ELISA). It may be used as
CC part of a method that allows determination of individual Listeria
CC species, esp. L. monocytogenes.
XX
XX
SQ Sequence 9 AA:
Query Match 100.0%; Score 9; DB 14; Length 9;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSPVPAPTQ 9
Db 1 vstpvaptq 9
RESULT 2
ID AAR73891 standard; peptide: 21 AA.
XX
XX AAR73891:
XX
XX
XX 05-DEC-1995 (first entry)
XX
XX
XX Listeria monocytogenes protein P60 precursor peptide 144-164.
XX
XX
XX Listeria monocytogenes; protein P60 precursor; vaccine;
KW meningitis related homologous antigenic sequence; MRHAS; RV-1;
KW immunoassay; diagnosis; treatment; prophylactic; bacterial;
KW viral; peptide 144-164.
XX
XX
XX Listeria monocytogenes.
OS
XX
XX WO9509232-A.
XX
XX
XX 06-APR-1995.
XX
XX
XX 28-SEP-1994; 94WO-CA00516.
XX
XX
XX 28-SEP-1993; 93US-0127499.
XX
XX
XX (SHAR/) SHARMA L R.
XX
XX (VALS/) VAN ALSTYNE D.
XX
XX
XX Sharma LR, Van Alstyne D;
XX
XX
XX WPI; 1995-147431/19.
XX
XX
XX New peptide(s) and corresp. antibodies for the treatment of
PT meningitis - the peptide(s) corresp. to homologous antigenic
PT sites on bacterial and viral agents and on chemokine(s), used for
PT detecting and preventing meningitis
XX
XX
XX Claim 34; Page 74; 98pp; English.
XX
XX
XX AAR7913 is the Listeria monocytogenes protein P60 precursor. It
CC contains the meningitis related antigenic sequences (MRHAS) claimed
CC in AAR73891-R73894 and AAR73903-R73906, which are recognised by a
CC monoclonal antibody from the hybridoma Rubella virus (RV)-1. The
CC claimed MRHAS peptides may be used in immunoassays to diagnose the
CC presence of bacterial and/or viral meningitis agents in a sample,
CC or in prophylactic and therapeutic meningitis treatments. The
CC peptides may also be used as vaccines against meningitis.

```

XX Sequence 21 AA;

Query Match 100.0%; Score 9; DB 16; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPPTQ 9  
| | | | | | | | | |  
DB 2 vstpvapptq 10

# RESULT 3

AAR45178 ID AAR45178 standard; Protein; 478 AA.

AC AAR45178;

DT 16-JUN-1994 (first entry)

DE Listeria p60 protein.

KW Immunogenic polypeptide; antibodies; immunoassay; conjugate.

OS Listeria monocytogenes.

PN DE4318450-A.

PD 16-DEC-1993.

PF 03-JUN-1993; 93DE-4318450.

PR 11-JUN-1992; 92DE-4219111.

PR 25-NOV-1992; 92DE-4239567.

PA (MERE) MERCK PATENT GMBH.

PI Schubert P, Neumann S, Pawelzik W, Linxweiler W, Burger C;

PI Hofmann G, Bubert A, Goedel W, Koehler S;

DR WPI; 1993-406956/51.

PT New primers for PCR detection of Listeria - including individual

PT species, also new peptide(s) for raising antibodies for

PT immunochemical detection

PS Disclosure: Fig 4; 19pp; German.

CC The sequence is that of the Listeria p60 protein. Antibodies

CC generated against the protein can be used in the detection

CC of Listeria by immunoassay (partic. ELISA). The detection method

CC allows determination of individual Listeria species, esp.

CC L. monocytogenes.

SO Sequence 478 AA;

# RESULT 4

AAR73913 ID AAR73913 standard; Protein; 484 AA.

AC AAR73913;

DT 05-DEC-1995 (first entry)

DE Listeria monocytogenes protein p60 precursor.

XX Listeria monocytogenes; protein p60 precursor; vaccine;

KW meningitis related homologous antigenic sequence; MRHAS; RV-1;

KW immunoassay; diagnosis; treatment; prophylactic; bacterial;

XX viral.

OS Listeria monocytogenes.

PN WO9509232-A.

PD 06-APR-1995.

PF 28-SEP-1994; 94WO-CA00516.

PR 28-SEP-1993; 93US-0127499.

PA (SHAR/) SHARMA L R.

PA (VALS/) VAN ALSTYNE D.

PI Sharma LR, Van Alstyne D;

DR WPI; 1995-147431/19.

PT New peptide(s) and corresp. antibodies for the treatment of

PT meningitis - the peptide(s) corresp. to homologous antigenic

PT sites on bacterial and viral agents and on chemokine(s), used for

PT detecting and preventing meningitis

PS Claim 47; Fig 7/10; 98pp; English.

CC AAR73913 is the Listeria monocytogenes protein p60 precursor. It

CC contains the meningitis related antigenic sequences (MRHAS) claimed

CC in AAR73891-R73894 and AAR73903-R73906, which are recognised by a

CC monoclonal antibody from the hybridoma Rubella virus (RV)-1. The

CC claimed MRHAS peptides may be used in immunoassays to diagnose the

CC presence of bacterial and/or viral meningitis agents in a sample,

CC or in prophylactic and therapeutic meningitis treatments. The

CC peptides may also be used as vaccines against meningitis.

CC NB: Identified by matching corresponding MRHAS peptides.

SO Sequence 484 AA;

OY 1 VSTPVAPPTQ 9

DB 145 vstpvapptq 153

# RESULT 5

AAR45172 ID AAR45172 standard; Protein; 12 AA.

AC AAR45172;

DT 16-JUN-1994 (first entry)

DE Listeria p60 peptide epitope.

KW Listeria monocytogenes; antibodies; immunoassay; conjugate.

OS Synthetic.

PN DE4318450-A.

PD 16-DEC-1993.

PF 03-JUN-1993: 93DE-4318450.  
 XX 11-JUN-1992: 92DE-4219111.  
 PR 25-NOV-1992: 92DE-4239567.  
 XX  
 PA (MERE ) MERCK PATENT GMBH.  
 XX  
 PI Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;  
 PI Hofmann G, Bubert A, Goebel W, Koehler S;  
 XX  
 DR WPI: 1993-406956/51.  
 XX  
 PT New primers for PCR detection of *Listeria* - including individual  
 PT species, also new peptide(s) for raising antibodies for  
 PT immunochemical detection  
 XX  
 PS Disclosure: Fig 2; 19pp; German.  
 XX  
 CC The sequence is that of a *Listeria* p60 peptide epitope which  
 CC which may be used in the prodn. of antibodies for the detection  
 CC of *Listeria* by immunoassay (partic. ELISA). It may be used as  
 CC part of a method that allows determination of individual *Listeria*  
 CC species, esp. *L. monocytogenes*.  
 CC  
 SQ Sequence 12 AA;

Query Match 88.9%; Score 8; DB 14; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.017;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 STPVAPTO 9  
 |||||  
 Db 1 stpvaptg 8

RESULT 6  
 AAR54615  
 ID AAR54615 standard; Protein: 6 AA.  
 XX  
 AC AAR54615;  
 XX  
 DT 16-JUN-1994 (first entry)  
 XX  
 DE *Listeria* p60 peptide epitope.  
 XX  
 KW *Listeria* monocytogenes; antibodies; immunoassay; conjugate.  
 XX  
 OS Synthetic.  
 XX  
 PN DE4318450-A.  
 XX  
 PD 16-DEC-1993.  
 XX  
 PF 03-JUN-1993: 93DE-4318450.  
 XX  
 PR 11-JUN-1992: 92DE-4219111.  
 PR 25-NOV-1992: 92DE-4239567.  
 XX  
 PA (MERE ) MERCK PATENT GMBH.  
 XX  
 PI Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;  
 PI Hofmann G, Bubert A, Goebel W, Koehler S;  
 XX  
 DR WPI: 1993-406956/51.  
 XX  
 PT New primers for PCR detection of *Listeria* - including individual  
 PT species, also new peptide(s) for raising antibodies for  
 PT immunochemical detection  
 XX  
 PS Disclosure: Page 3; 19pp; German.  
 XX  
 CC The sequence is that of a *Listeria* p60 peptide epitope which

CC which may be used in the prodn. of antibodies for the detection  
 CC of *Listeria* by immunoassay (partic. ELISA). It may be used as  
 CC part of a method that allows determination of individual *Listeria*  
 CC species, esp. *L. monocytogenes*.  
 CC  
 SQ Sequence 6 AA;

Query Match 66.7%; Score 6; DB 14; Length 6;  
 Best Local Similarity 100.0%; Pred. No. 3.4e+05;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPTO 9  
 |||||  
 Db 1 pvaptg 6

RESULT 7  
 AAR45173  
 ID AAR45173 standard; Protein: 10 AA.  
 XX  
 AC AAR45173;  
 XX  
 DT 16-JUN-1994 (first entry)  
 XX  
 DE *Listeria* p60 peptide epitope.  
 XX  
 KW *Listeria* monocytogenes; antibodies; immunoassay; conjugate.  
 XX  
 OS Synthetic.  
 XX  
 PN DE4318450-A.  
 XX  
 PD 16-DEC-1993.  
 XX  
 PF 03-JUN-1993: 93DE-4318450.  
 XX  
 PR 11-JUN-1992: 92DE-4219111.  
 PR 25-NOV-1992: 92DE-4239567.  
 XX  
 PA (MERE ) MERCK PATENT GMBH.  
 XX  
 PI Schubert P, Neumann S, Pawelzik M, Linxweiler W, Burger C;  
 PI Hofmann G, Bubert A, Goebel W, Koehler S;  
 XX  
 DR WPI: 1993-406956/51.  
 XX  
 PT New primers for PCR detection of *Listeria* - including individual  
 PT species, also new peptide(s) for raising antibodies for  
 PT immunochemical detection  
 XX  
 PS Disclosure: Fig 2; 19pp; German.  
 XX  
 CC The sequence is that of a *Listeria* p60 peptide epitope which  
 CC which may be used in the prodn. of antibodies for the detection  
 CC of *Listeria* by immunoassay (partic. ELISA). It may be used as  
 CC part of a method that allows determination of individual *Listeria*  
 CC species, esp. *L. monocytogenes*.  
 CC  
 SQ Sequence 10 AA;

Query Match 66.7%; Score 6; DB 14; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPTO 9  
 |||||  
 Db 1 pvaptg 6

RESULT 8  
 AAR45159

```

ID  AAR45159 standard; Protein: 20 AA.
XX
AC  AAR45159;
XX
DT  16-JUN-1994 (first entry)
XX
DE  Listeria p60 peptide epitope.
XX
KW  Listeria monocytogenes; antibodies; immunoassay; conjugate.
XX
OS  Synthetic.
XX
FH  Key
FT  Region
FT  Location/Qualifiers
FT  1..7
FT  /note= "opt. 0-7 amino acid residues"
FT  14..20
FT  /note= "opt. 0-7 amino acid residues"
XX
DE  DE4318450-A.
XX
PD  16-DEC-1993.
XX
PE  03-JUN-1993; 93DE-4318450.
XX
PR  11-JUN-1992; 92DE-4219111.
XX  25-NOV-1992; 92DE-4239567.
XX
PA  (MERCK ) MERCK PATENT GMBH.
XX
PI  Schubert P, Neumann S, Pawelzik M, Linxweller W, Burger C;
XX  Hofmann G, Bubert A, Goebel W, Koehler S;
XX  WPI: 1993-406956/51.
XX
DR  New primers for PCR detection of Listeria - including individual
PT  species, also new peptide(s) for raising antibodies for
PT  immunochemical detection
XX
PS  Claim 3; Page 4; 19pp; German.
XX
CC  The sequence is that of a Listeria p60 peptide epitope which
CC  which may be used in the prodn. of antibodies for the detection
CC  of Listeria by immunoassay (partic. ELISA). It may be used as
CC  part of a method that allows determination of individual Listeria
CC  species, esp. L. monocytogenes.
XX
SQ  Sequence 20 AA:

Query Match 66.7%; Score 6; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  4 PVAPRO 9
    |||||
DB  8 pvaprc 13

RESULT 9
ID  AAY37718 standard; Protein: 157 AA.
XX
AC  AAY37718;
XX
DT  07-OCT-1999 (first entry)
XX
DE  Amino acid sequence of a Chlamydia trachomatis protein.
XX
KW  Vaccine; eye disease; conventional trachoma; nongonococcal urethritis;
KW  paratrachoma; inclusion conjunctivitis; genital disease; perithea-
KW  nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW  bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX

```

```

OS  Chlamydia trachomatis.
XX
PN  W09928475-A2.
XX
PD  10-JUN-1999.
XX
PE  27-NOV-1998; 98WO-1B01939.
XX
PR  04-NOV-1998; 98US-0107077.
XX  28-NOV-1997; 97ER-0015041.
XX  17-DEC-1997; 97ER-0016034.
XX
PA  (GIST ) GENSET.
XX
PI  Griffais R;
XX
DR  WPI: 1999-371125/31.
XX
PT  Genome sequence of Chlamydia trachomatis
XX
PS  Disclosure: Page 1323-1324; 1755pp; English.
XX
CC  AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC  of Chlamydia trachomatis (see A4201425). The polypeptides can be used as
CC  vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC  can also be used to control growth of the microorganism. Chlamydia
CC  trachomatis is responsible for a large number of diseases, e.g. eye
CC  diseases such as conventional trachoma, nongonococcal urethritis,
CC  paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC  nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
CC  perithea-
CC  and venereal lymphogranulomatosis. The polypeptides of the invention
CC  may be of use in treating these diseases.
XX
SQ  Sequence 157 AA:

Query Match 66.7%; Score 6; DB 20; Length 157;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  2 SRVPAP 7
    |||||
DB  14 srvpap 19

RESULT 10
ID  AAB41666 standard; Protein: 343 AA.
XX
AC  AAB41666;
XX
DT  08-FEB-2001 (first entry)
XX
DE  Human ORFX ORF1430 polypeptide sequence SEQ ID NO:2860.
XX
KW  Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW  vulnereary; antiproliferative; antiparkinsonian; neurotrophic; neuroprotective;
KW  anticonvulsant; osteoprotic; antiarthritic; immunosuppressant; cardiant;
KW  immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW  hypotensive; dermatologic; immunosuppressive; antiinflammatory;
KW  antiviral; antibacterial; antifungal; antineumatic; antithyroid;
KW  antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW  neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW  cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW  cholesterol ester storage; systemic lupus erythematosus; infection;
KW  severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW  allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW  bone damage; cartilage damage; antinflammatory disease; coagulation;
KW  thrombosis; contraceptive.
XX
OS  Homo sapiens.
XX

```

PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000WO-US08621.  
 XX  
 PR 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkets RA, Leach M;  
 XX  
 DR WPI; 2000-602362/57.  
 DR N-PSDB; AAC75875.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 PS Claim 11; Page 2100-2101; 5507pp; English.  
 XX  
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;  
 CC antiparasitic; antiparkinsonian; nootropic; neuroprotective;  
 CC osteopathic; anticoagulant; antiarthritic; immunosuppressant;  
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antiinflammatory; antibacterial; antiviral; antifungal; antineumatic;  
 CC antihydroid; and antianemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy.  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 CC  
 SQ Sequence 343 AA;  
 Query Match 66.7%; Score 6; DB 21; Length 343;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 3 TPVAPT 8  
 | | | | |  
 Db 289 tpvapt 294  
 RESULT 11  
 ID AAY05477 standard; Protein; 750 AA.  
 AC AAY05477;  
 XX  
 DT 07-JUL-1999 (first entry)  
 XX  
 DE C. albicans Rbt1 protein sequence.  
 XX  
 XX Rbt1; detection; TUP1 gene function; pathogenic potential; vaccine;  
 KM virulence inhibitor; infection; diagnosis; filamentous growth;  
 XX antibiotic resistance.  
 OS Candida albicans.  
 PN

XX  
 PN WO918115-A1.  
 XX  
 PD 15-APR-1999.  
 XX  
 PF 02-OCT-1998; 98WO-US20655.  
 XX  
 PR 01-OCT-1998; 98US-0165239.  
 PR 02-OCT-1997; 97US-0061058.  
 PR 18-DEC-1997; 97US-0068065.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Braun B, Johnson AD;  
 XX  
 DR WPI; 1999-263994/22.  
 DR N-PSDB; AAX36552.  
 XX  
 PT Nucleic acid encoding Rbt1 polypeptides of Candida albicans  
 XX  
 PS Claim 28; Fig 3b; 82pp; English.  
 XX  
 CC This sequence is the Candida albicans Rbt1 protein of the invention.  
 CC The Rbt1 gene and its fragments are useful for detecting or quantifying  
 CC C. albicans nucleic acid in hybridization or amplification methods; as  
 CC therapeutic agents; as indicators of TUP1 gene function (and thus of  
 CC pathogenic potential); as a polypeptide processing signal or template for  
 CC various functional domains of Rbt1. Plasmids containing the Rbt1 gene can  
 CC be used as vaccines. Rbt1 is used to detect or quantify specific  
 CC antibodies against Rbt1, while antibodies against Rbt1 can be used to  
 CC detect Rbt1 in standard immunoassays. Rbt1 can also be used for drug  
 CC screening, rational drug design and identification of proteins that  
 CC interact physically with Rbt1. Cells that express Rbt1 can be used to  
 CC screen for agents that control virulence in C. albicans. Antibodies  
 CC against Rbt1 may be used to purify Rbt1 and to raise anti-idiotypic  
 CC antibodies. Modulation of Rbt1 function, specifically by disrupting the  
 CC corresponding gene, is used to inhibit virulence of C. albicans. The Rbt1  
 CC gene and protein allow diagnosis of infection without the need for  
 CC culturing and microscopic examination. The presence of Rbt1 indicates  
 CC lack of TUP1 activity and/or filamentous growth, and also the degree of  
 CC virulence and/or resistance/susceptibility to particular antibiotics, for  
 CC prognosis or selection of appropriate treatments.  
 CC  
 SQ Sequence 750 AA;  
 Query Match 66.7%; Score 6; DB 20; Length 750;  
 Best Local Similarity 100.0%; Pred. No. 99;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2 STPVAP 7  
 | | | | |  
 Db 650 stpvap 655  
 RESULT 12  
 ID AAB59817 standard; Protein; 999 AA.  
 AC AAB59817;  
 XX  
 DT 04-APR-2001 (first entry)  
 XX  
 DE TutD protein #8.  
 XX  
 KM Toluene degradation; enzyme; waste degradation; TutD.  
 XX  
 XX Thaueria aromatica.  
 OS Xanthomonas maltophilia.  
 OS Geobacter metallireducens.  
 OS Azorocus toluilyticus.  
 XX  
 PN WO200072650-A2.



XX 07-DEC-2000.  
PD  
XX 24-MAY-2000; 2000WO-US14298.  
PE  
XX 01-JUN-1999; 99US-0323872.  
PR  
XX (UYOH-) UNIV OHIO.  
PA  
XX Coschigano PW;  
PI  
XX WPI: 2001-041080/05.  
DR  
XX N-PSDB: AAF23625, AAF23627.  
XX  
PT Composition comprising toluene degrading enzyme useful for biological  
PT treatment of organic compounds, especially for degrading toluene or its  
PT analogs -  
PS  
XX Disclosure: Fig 5; 122pp; English.  
PS  
XX The present invention relates to toluene degrading enzyme genes and  
CC proteins tuth (see AAF23629 and AAB59831), turt (AAF23630 and AAB59832),  
CC turt (AAF23631 and AAB59833) and turt (AAF23632 and AAB59834). The  
CC toluene degrading enzymes are homologues of pyruvate formate lyase. The  
CC toluene degrading enzymes are useful for biological treatment of organic  
CC compounds and in particular for the degradation of toluene and its  
CC analogs contained in liquid or solid waste source. The present sequence  
CC is a protein sequence for toluene degrading enzyme, turt.  
CC  
XX Sequence 999 AA;  
SQ

Query Match 66.7%; Score 6; DB 22; Length 999;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTPVA 6  
|||||  
Db 757 vstpva 762

RESULT 13  
AAB59827  
ID AAB59827 standard; Protein: 1592 AA.  
XX  
AC AAB59827;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Protein #4 encoded by turtD/E gene.  
XX  
KW Toluene degradation; enzyme: waste degradation; Tute; TurtD.  
XX  
OS Thauera aromatica.  
OS Xanthomonas maltophilia.  
OS Geobacter metallireducens.  
OS Azorarcus toluyticus.  
XX  
XX WO200072650-A2.  
PN  
XX  
XX 07-DEC-2000.  
PD  
XX 24-MAY-2000; 2000WO-US14298.  
PE  
XX 01-JUN-1999; 99US-0323872.  
PR  
XX (UYOH-) UNIV OHIO.  
PA  
XX  
PI Coschigano PW;  
XX  
DR WPI: 2001-041080/05.  
DR N-PSDB: AAF23627.  
XX

PT Composition comprising toluene degrading enzyme useful for biological  
PT treatment of organic compounds, especially for degrading toluene or its  
PT analogs -  
PS  
XX Disclosure: Fig 12; 122pp; English.  
PS  
XX The present invention relates to toluene degrading enzyme genes and  
CC proteins tuth (see AAF23629 and AAB59831), turt (AAF23630 and AAB59832),  
CC turt (AAF23631 and AAB59833) and turt (AAF23632 and AAB59834). The  
CC toluene degrading enzymes are homologues of pyruvate formate lyase. The  
CC toluene degrading enzymes are useful for biological treatment of organic  
CC compounds and in particular for the degradation of toluene and its  
CC analogs contained in liquid or solid waste source. The present sequence  
CC is a protein sequence encoded by toluene degrading enzyme gene, turtD/E.  
CC  
XX Sequence 1592 AA;  
SQ

Query Match 66.7%; Score 6; DB 22; Length 1592;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTPVA 6  
|||||  
Db 1350 vstpva 1355

RESULT 14  
AAV21365  
ID AAV21365 standard; Protein: 7 AA.  
XX  
AC AAV21365;  
XX  
DT 22-JUL-1999 (first entry)  
XX  
DE Human HUPF-I mutant protein fragment 17.  
XX  
KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;  
KW frameshift mutation; age-related disease; neurodegenerative disorder;  
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;  
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;  
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;  
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;  
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;  
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;  
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMG-C; NSP-A;  
KW high mobility group protein-C; neuroendocrine specific protein A.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
XX WO9845322-A2.  
PN  
XX  
XX 15-OCT-1998.  
PD  
XX 02-APR-1998; 98WO-IB00705.  
PE  
XX 10-APR-1997; 97US-0043163.  
PR  
XX (UYUT-) RIJKSUNIV UTRECHT.  
PA (ROYA-) ROYAL NETHERLANDS ACADEMY OF SCI.  
PA (UYRO-) UNIV ROTTERDAM ERASMUS.  
XX  
PI Burbach JPH, Grosveld FG, Van Leeuwen FW;  
XX  
XX WPI: 1998-609901/51.  
DR N-PSDB: AAX75768.  
XX  
PT Diagnosing disease by detecting frameshift mutations in RNA or  
PT corresponding protein mutations - used to diagnose cancer and  
PT neurological diseases, particularly Alzheimer's disease, and also  
PT for treatment and prevention with specific ribozymes or wild-type  
PT RNA

XX This invention describes a novel method for the diagnosis of a disease  
PS disclosure: Figure 17; 258pp; English.  
XX  
CC  
CC  
CC caused by, or associated with, an RNA molecule that has a frameshift  
CC mutation. The method is used to diagnose age-related diseases, especially  
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's  
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,  
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II  
CC and many others listed) or susceptibility to these disorders. The method  
CC allows a definitive diagnosis of Alzheimer's disease in living patients,  
CC at an early stage. It is based on the observation that disease may be  
CC caused by mutations in RNA rather than DNA. The invention describes the  
CC use of neuronal system RNA molecules, specifically proteins including  
CC beta-amyloid precursor protein (beta-APP), the microtubule associated  
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule  
CC associated protein 2 (MAP2), neurofilament-F, neurofilament-M,  
CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic  
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma  
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group  
CC protein-C (HMGp-C) and neuroendocrine specific protein A.  
XX  
XX  
XX Sequence 7 AA;

Query Match	55.6%	Score 5;	DB 19;	Length 7;
Best Local Similarity	100.0%	Pred. No.	3.4e+05;	
Matches	5;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0

QY	4	PVAPT	8
Db	2	pvapt	6

## RESULT 15

ID AAW70120 standard; peptide; 9 AA.

AC AAW70120;

DT 07-DEC-1998 (first entry)

DE Peptide produced by subtilisin digestion of 45 kDa MPT 32 glycoprotein.

KW Glycopeptide; subtl11sn; reversed-phase HPLC; Edman degradation; FAB-MS;

KW antigen.

OS Mycobacterium tuberculosis.

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

/note= "O-glycosylated with an alpha-mannose residue"

PN WO9829132-A1

PD 09-JUL-1998

PF 29-DEC-1997; 97WO-US24189.

PR 31-DEC-1996; 96US-0034003.

PA (UYNV ) UNIV NEW YORK STATE.

Belisle JT, Laal S, Zolla-Pazner S;

DR WPI; 1998-387787/33.

PT Early detection of mycobacterial infection - by testing a biological  
PT fluid sample from a subject for the presence of antibodies reactive  
PT with *Mycobacterium tuberculosis* antigens

PS Example 3; Page 71; 170pp; English.

XX Sequences AAM70109, AAM70115-w70141 are N-terminal amino acid sequences of glycopeptides, or non-glycosylated peptides generated by subtilisin digestion of the 45 kDa MPr 32 glycoprotein. These sequences were obtained by purifying the 45 kDa protein, digesting it with subtilisin, and then separating the products by reversed-phase HPLC on a column. CC This digestion yielded 26 peptides, of which five were found to be CC glycosylated (AAM70109, AAM70115, AAM70120, AAM70126, and AAM70128). All CC of the peptides were subjected to both Fast Atom Bombardment-Mass CC Spectrometry (FAB-MS), and automated Edman degradation to determine their CC molecular weight and N-terminal amino acid sequence, respectively. The CC discovery of the complete primary structure of MPr 32 glycoprotein CC enables further work to be performed and the possibility of it being used CC as an antigen for the early detection of mycobacterial disease.

Query Match	55.6%	Score 5;	DB 19;	Length 9;
Best Local	Similarity	Pred. No.	3.4e+05;	
Matches	5;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0

Qy	3	TPVAP	7
Db	1	tpvap	5

## RESULT 16

ID AAW70126 standard; peptide; 11 AA

AC AAW70126;

DT 07-DEC-1998 (first entry)

DE Peptide produced by subtilisin digestion of 45 kDa MPT 32 glycoprotein.

KW Glycopeptide; subtilisin; reversed-phase HPLC; Edman degradation; FAB-MS;

KW antligen.

Mycobacterium tuberculosis.

EH	Key	Location/Qualifiers
EH	Key	Location/Qualifiers

FT /note= "O-glycosylated with an alpha-mannose residue"

PN WO9829132-A1

PD 09-JUL-1998

PF 29-DEC-1997; 97WO-US24189.

PR 31-DEC-1996; 96US-0034003.

PA (UYN Y ) UNIV NEW YORK STATE.

PI Belisle JT, Laal S, Zollla-pazner S;

DR WPI; 1998-387787/33

PT Early detection of mycobacterial infection - by testing a biological  
PT fluid sample from a subject for the presence of antibodies reactive  
PT with *Mycobacterium tuberculosis* antigens

PS Example 3; Page 71; 170pp; English.

CC Sequences AAM70109, AAM70115-W70141 are N-terminal amino acid sequences  
CC of glycopeptides, or non-glycosylated peptides generated by subtilisin  
CC digestion of the 45 kDa MPr 32 glycoprotein. These sequences were  
CC obtained by purifying the 45 kDa protein, digesting it with subtilisin,  
CC and then separating the products by reversed-phase HPLC on a column.  
CC This digestion yielded 26 peptides, of which five were found to be

CC glycosylated (AAW70109, AAW70115, AAW70120, AAW70126, and AAW70128). All  
 CC of the peptides were subjected to both Fast Atom Bombardment-Mass  
 CC Spectrometry (FAB-MS), and automated Edman degradation to determine their  
 CC molecular weight and N-terminal amino acid sequence, respectively. The  
 CC discovery of the complete primary structure of MPT 32 glycoprotein  
 CC enables further work to be performed and the possibility of it being used  
 CC as an antigen for the early detection of mycobacterial disease.

SQ Sequence 11 AA:

Query Match 55.6%; Score 5; DB 19; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 pvpap 7  
 11111  
 Db 1 tpvap 5

RESULT 17

AAW72848  
 ID AAW72848 standard; Protein; 11 AA.

AC AAW72848;

DT 09-MAY-2001 (first entry)

DE Human p53 A76T mutated peptide.

XX Human; p53 mutation; mutant; muten; gene therapy; supertransactivating;  
 KM cancer.

XX Homo sapiens.

OS Synthetic.

PN WO200109325-A2.

PD 08-FEB-2001.

PF 28-JUL-2000; 2000WO-US20538.

PR 30-JUL-1999; 99US-0146634.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Resnick MA, Inga A;

DR WPI; 2001-123321/13.

XX New isolated mutated human p53 polypeptides for inducing toxicity in a  
 PT cell, treating cancer and identifying compounds that mimic toxic or  
 PT supertransactivating mutations -  
 PS Claim 6; Page -; 144pp; English.

CC The present invention provides a number of peptides derived from the  
 CC human p53 protein and containing specific mutations. These are useful for  
 CC isolating supertransactivating and toxic p53 mutants, and for inhibiting  
 CC cell growth and treating cancer. The present sequence is a peptide  
 CC derived from the wild-type human p53 protein shown in AAW72878.  
 CC Note: The present sequence is not shown in the specification but is  
 CC derived from that given in Genbank acc. no. X02469 (see AAW72878).

SQ Sequence 11 AA:

Query Match 55.6%; Score 5; DB 22; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
 11111

Db 3 pvpap 7

RESULT 18

AAW82295  
 ID AAW82295 standard; Peptide; 20 AA.

AC AAW82295;

DT 15-MAR-1999 (first entry)

DE Dynamin 41.20 zelan120 peptide.

XX Targeting agent; gastrointestinal tract; brain; drug delivery;  
 KM drug targeting; phage display; dynamin II; human.

XX Homo sapiens.

OS Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1 /note="N-terminal dansyl group"

PN WO9851825-A1.

PD 19-NOV-1998.

PF 15-MAY-1998; 98WO-US10079.

PR 15-MAY-1997; 97US-0857046.

PA (CYNO-) CYTOGEN CORP.

PA (ELAN-) ELAN CORP PLC.

PI Alvarez VL, O'Mahony DJ, Seveso M;

DR WPI; 1999-009778/01.

XX New targeting agents enabling transport of active agents through  
 PT human/animal tissue, or their uptake - useful in the treatment and  
 PT prevention of diseases, especially of the gastro-intestinal tract  
 PT (GIT)

PS Example 4; Page 53; 83pp; English.

XX This dansylated peptide, termed dynamin 41.20 zelan120 peptide,  
 CC is a truncated fragment of dynamin 41.2 zelan086 peptide (see  
 CC AAW82294). Dynamin 41.2 is derived from human dynamin II and shows  
 CC homology to peptide 41.1 (see AAW82293), which was isolated from a  
 CC phage display library and shown to have the ability to enhance  
 CC transport of active agents from the gastrointestinal tract (GIT) to  
 CC brain. Dynamin 41.2 and peptide fragments of it (see AAW82295-97)  
 CC were compared with peptide 41.1 for their ability to enhance  
 CC transport of PLGA particles in vivo and in vitro. The invention  
 CC relates to targeting agents capable of permitting or facilitating  
 CC transport of an active agent through human or animal GIT tissue.  
 CC The targeting agents are peptides (see AAW82286-93), or their  
 CC derivatives (e.g. fragments, see AAW82300-10) and peptidomimetics,  
 CC and the nucleotide sequences coding for them (see AAV73312-19). They  
 CC have use in facilitating transport of active agents from the GIT  
 CC lumen into the systemic blood system and/or in targeting active  
 CC agents to the GIT. Preferably, the active agent is a drug or  
 CC drug-containing nanoparticle or microparticle.

SQ Sequence 20 AA:

Query Match 55.6%; Score 5; DB 20; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 11111

Db 5 vslpv 9

RESULT 19  
AAW82310  
ID AAW82310 standard; Peptide: 23 AA.

AC AAW82310;  
XX  
XX 15-MAR-1999 (first entry)  
XX  
DE Transport peptide fragment.  
XX  
XX Targeting agent; gastrointestinal tract; brain; drug delivery;  
KW drug targeting; dynamint; human; transport peptide.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
PN WO9851825-A1.  
XX  
PD 19-NOV-1998.  
XX  
PF 15-MAY-1998: 98WO-US10079.  
XX  
PR 15-MAY-1997: 97US-0857046.  
XX  
PA (CYTO-) CYTOGEN CORP.  
PA (ELAN-) ELAN CORP PLC.  
XX  
XX Alvarez VL, O'Mahony DJ, Seveso M;  
PI WPI: 1999-009778/01.  
XX  
XX New targeting agents enabling transport of active agents through  
PT human/animal tissue, or their uptake - useful in the treatment and  
PT prevention of diseases, especially of the gastro-intestinal tract  
XX (GIT)  
XX  
PS Claim 10; Page 66; 83pp; English.  
XX  
XX This is a fragment of dynamin 41.2 (see AAW82294), a peptide derived  
CC from human dynamin II. Claimed transport or uptake targeting  
CC agents can include this, or other, claimed peptide fragments (see  
CC AAW82300-10). Such transport or targeting agents permit or  
CC facilitate the transport of an active agent through a human or  
CC animal tissue such as gastrointestinal tract lumen, or uptake of  
CC the active agent into the human or animal tissue. The active agent  
CC is preferably a drug or drug-containing nanoparticle or  
CC microparticle. Therapeutic methods of administration,  
CC pharmaceutical compositions and formulations based of the  
CC targeting peptides are also provided.  
XX  
XX  
SQ Sequence 23 AA:

Query Match 55.6%; Score 5; DB 20; Length 23;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
XX | | | | |  
XX 5 vslpv 9

Db 5 vslpv 9

RESULT 20  
AAW79982  
ID AAW79982 standard; peptide: 25 AA.  
XX  
XX AAW79982;  
XX  
XX 01-MAY-1996 (first entry)  
XX

DE Signal transduction regulatory tyrosine kinase SH3 binding peptide.  
XX  
XX Signal transduction; src-family thymidine kinase; lipid kinase;  
KW tyrosine kinase; adaptor; inhibition; regulation; immunodeficiency;  
KW allergy; inflammation; cancer; autoimmune disease; Epstein-Barr virus;  
KW immunoproliferative disease; bovine leukaemia virus; EBV; BLV;  
KW chronic fatigue syndrome; infectious mononucleosis.  
XX  
XX Homo sapiens.  
OS  
XX  
PN WO9524915-A1.  
XX  
PD 21-SEP-1995.  
XX  
PF 17-MAR-1995: 95WO-US03438.  
XX  
PR 17-MAR-1994: 94US-0215116.  
XX  
XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.  
PA Cambler JC, Clark MR, Pleiman CM;  
PI WPI: 1995-336809/43.  
XX  
XX Regulation of signal transduction pathways, esp. involving  
PT src-family thymidine kinase - useful for treatment of allergic and  
PT inflammatory responses  
XX  
XX  
PS Claim 10; Page 101; 131pp; English.  
XX  
XX AAW79982-83 are signal transduction regulatory ARH1 peptides. They  
CC bind the SH3 domain of src family tyrosine kinases. ARH1 regulatory  
CC peptides are capable of regulating the activity of tyrosine kinases,  
CC lipid kinases, thymidine kinases and adaptor molecules. They can be  
CC used in a method to regulate signal transduction in such cells as  
CC B cells, T cells, macrophages, dendritic cells and pluripotent stem  
CC cells. The method allows regulation and treatment of allergic and  
CC inflammatory responses, autoimmune disease, immunodeficiency and  
CC immunoproliferative diseases, cancer and disease caused by  
CC Epstein-Barr and bovine leukaemia viruses.  
XX  
XX  
SQ Sequence 25 AA:

Query Match 55.6%; Score 5; DB 16; Length 25;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
XX | | | | |  
XX 16 tpvap 20

Db 16 tpvap 20

RESULT 21  
AAW78130  
ID AAW78130 standard; Protein: 27 AA.  
XX  
XX AAW78130;  
XX  
XX 13-APR-1999 (first entry)  
XX  
XX  
XX Human secreted protein encoded by gene 5 clone HPBD36.  
DE  
XX  
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX  
XX Homo sapiens.  
OS

XX Key Location/Qualifiers  
 FH Misc-difference 27  
 FT /label= unknown  
 XX  
 XX MO9856804-A1.  
 PD 17-DEC-1998.  
 XX  
 XX 11-JUN-1998; 98WO-US12125.  
 XX 02-OCT-1997; 97US-0061060.  
 PR 13-JUN-1997; 97US-0049547.  
 PR 13-JUN-1997; 97US-0049548.  
 PR 13-JUN-1997; 97US-0049549.  
 PR 13-JUN-1997; 97US-0049550.  
 PR 13-JUN-1997; 97US-0049606.  
 PR 13-JUN-1997; 97US-0049607.  
 PR 13-JUN-1997; 97US-0049608.  
 PR 13-JUN-1997; 97US-0049609.  
 PR 13-JUN-1997; 97US-0049610.  
 PR 13-JUN-1997; 97US-0049611.  
 PR 13-JUN-1997; 97US-0050566.  
 PR 13-JUN-1997; 97US-0050901.  
 PR 13-JUN-1997; 97US-0052989.  
 PR 18-AUG-1997; 97US-0051919.  
 PR 12-SEP-1997; 97US-0055984.  
 PR 12-SEP-1997; 97US-0058665.  
 PR 12-SEP-1997; 97US-0058668.  
 PR 12-SEP-1997; 97US-0058669.  
 PR 12-SEP-1997; 97US-0058750.  
 PR 12-SEP-1997; 97US-0058971.  
 PR 12-SEP-1997; 97US-0058972.  
 PR 12-SEP-1997; 97US-0060834.  
 PR 02-OCT-1997; 97US-0060841.  
 PR 02-OCT-1997; 97US-0060844.  
 PR 02-OCT-1997; 97US-0060865.  
 PR 02-OCT-1997; 97US-0061059.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Brewer LA, Ebner R, Ferrie AM, Feng P, Greene JM, Lafleur DW;  
 PI Moore PA, NI J, Olsen HS, Rosen CA, Ruben SM, Shi Y, Young P;  
 PI Yu GL;  
 XX WPI: 1999-080881/07.  
 DR N-PSDB; AAX04315.  
 XX  
 PT New isolated human genes and the secreted polypeptides they encode -  
 PT useful for diagnosis and treatment of e.g. cancers, neurological  
 PT disorders, immune diseases, inflammation or blood disorders  
 XX  
 XX Claim 11; Page 266; 380pp; English.  
 CC This sequence represents a secreted human protein encoded by the gene  
 CC clone detailed in the descriptor line. The gene can be used to generate  
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc  
 CC portion (e.g. AAX04302) for increasing the stability of the fused  
 CC protein as compared to the human protein only.  
 CC The invention relates to 86 novel genes and their fragments (nucleic  
 CC acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)  
 CC which are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also, pathological  
 CC conditions can be diagnosed by determining the amount of the new  
 CC polypeptides in a sample or by determining the presence of mutations in  
 CC the new polynucleotides. Specific uses are described for each of the 86  
 CC polynucleotides, based on which tissues they are most highly expressed in  
 CC (see AAX04311 for described uses).  
 XX  
 SQ Sequence 27 AA;

Query Match 55.6%; Score 5; DB 20; Length 27;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
 |||||  
 DB 9 pvapt 13

## RESULT 22

AAW82294  
 ID AAW82294 standard; Peptide; 43 AA.

AC AAW82294;

DT 15-MAR-1999 (first entry)

DE Dynamin 41.2 ZELAN086 peptide.

KW Targetting agent; gastrointestinal tract; brain; drug delivery;  
 KW drug targetting; phage display; dynamin II; human.

OS Homo sapiens.  
 OS Synthetic.

FT Key. Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal dansyl group"

XX MO9851825-A1.

PD 19-NOV-1998.

XX 15-MAY-1998; 98WO-US10079.

XX 15-MAY-1997; 97US-0857046.

XX (CYTO-) CYTOGEN CORP.  
 XX (ELAN-) ELAN CORP PLC.

XX Alvarez VL, O'Mahony DJ, Seveso M;

DR WPI: 1999-009778/01.

PT New targetting agents enabling transport of active agents through  
 PT human/animal tissue, or their uptake - useful in the treatment and  
 PT prevention of diseases, especially of the gastro-intestinal tract  
 PT (GIT)  
 XX  
 XX Example 4; Page 53; 83pp; English.

XX This dansylated peptide, termed dynamin 41.2 ZELAN086 peptide,  
 CC is derived from human dynamin II. It shows homology to peptide  
 CC 41.1 (see AAW82293), which was isolated from a phage display  
 CC library and shown to have the ability to enhance transport of  
 CC active agents from the gastrointestinal tract (GIT) to brain.  
 CC Dynamin 41.2 and peptide fragments of it (see AAW82295-97) were  
 CC compared with peptide 41.1 for their ability to enhance transport  
 CC of PLGA particles in vivo and in vitro. The invention relates to  
 CC targetting agents capable of permitting or facilitating transport  
 CC of an active agent through human or animal GIT tissue. The  
 CC targetting agents are peptides (see AAW82286-93), or their derivatives  
 CC (e.g. fragments, see AAW82300-10) and peptidomimetics, and the  
 CC nucleotide sequences coding for them (see AAV7312-19). They have  
 CC use in facilitating transport of active agents from the GIT lumen  
 CC into the systemic blood system and/or in targetting active agents  
 CC to the GIT. Preferably, the active agent is a drug or  
 CC drug-containing nanoparticle or microparticle.

XX Sequence 43 AA;

Query Match 55.6%; Score 5; DB 20; Length 43;

Best Local Similarity 100.0%; Pred. No. 71;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 VSRPV 5  
|||||  
DB 5 vsrpv 9

RESULT 23  
AAV01143  
ID AAV01143 standard; Protein; 45 AA.

AAV01143;

18-MAY-1999 (first entry)

Secreted protein encoded by gene 9 clone HSDY06.

Human; secreted protein; gene therapy; protein therapy; tissue; cancer;  
tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;  
developmental abnormality; foetal deficiency; Alzheimer's disease;  
cognitive disorder; schizophrenia; immunological disorder; mood disorder;  
immune deficiency disease; respiratory disorder; arthritis; skeletal;  
haematopoietic disorder; neural; osteoporosis; metabolic disorders;  
cardiovascular; endocrine; gastrointestinal; asthma; diagnosis.

Homo sapiens.

MO9901020-A2.

14-JAN-1999.

30-JUN-1998; 98MO-US13608.

12-SEP-1997; 97US-0058663.

01-JUL-1997; 97US-0051381.

01-JUL-1997; 97US-0051480.

12-SEP-1997; 97US-0058598.

(HUMA-) HUMAN GENOME SCI INC.

Carter KC, Endress GA, Feng P, Rosen CA, Ruben SM;

WPI; 1999-105663/09.

N-PSDB; AAX22119.

Claim 11; Page 146; 179pp; English.

The invention relates to nucleic acid sequences (AAX22111 to AAX22134)  
encoding human secreted proteins (AAV01135 to AAV01158). The secreted  
protein gene sequences are deposited with the ATCC under deposit number  
ATCC 209118. Host cells comprising recombinant vectors containing the  
nucleic acid sequences are used for the recombinant production of the  
secreted proteins. The polynucleotide and amino acid sequences are useful  
for are useful for preventing, treating or ameliorating medical  
conditions e.g. by protein or gene therapy. Pathological conditions can  
be also diagnosed by determining the amount of the new polypeptides in a  
sample or by determining the presence of mutations in the new  
polynucleotides. Specific uses are described for each of the  
polynucleotides, based on which tissues they are most highly expressed  
in, and include developing products for the diagnosis or treatment of  
cancer, tumours, developmental abnormalities and foetal deficiencies,  
autoimmune diseases, lymphomas, Alzheimer's and cognitive disorders,  
schizophrenia, immunological disorders, immune deficiency diseases  
(AIDS), mood disorders, respiratory disorders, arthritis, asthma,  
haematopoietic disorders, neural disorders, skeletal disorders,  
osteoporosis, metabolic disorders, cardiovascular disorders, endocrine  
disorders or gastrointestinal disorders. The polypeptides are also useful

CC for identifying their binding partners. The present sequence represents a  
CC human secreted protein (see descriptor line for gene number and clone  
CC identification).  
XX  
SQ Sequence 45 AA;

Query Match 55.6%; Score 5; DB 20; Length 45;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
|||||  
DB 33 tpvap 37

RESULT 24

AA33087  
ID AAB33087 standard; Protein; 48 AA.

AA33087;

25-JAN-2001 (first entry)

Pinus radiata transcription factor protein sequence #214.

Plant; transcription factor; gene expression; eucalyptus; pine; acacia;  
poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;  
basic helix-loop-helix zipper; homeotic; homeobox; MADS;  
homeodomain zipper; LIM domain; AP2; EREBS; zinc finger domain;  
type 2 Cys2His2; CCAAT box element; MYB.

Pinus radiata.

MO200053724-A2.

14-SEP-2000.

09-MAR-2000; 2000MO-US06112.

11-MAR-1999; 99US-0266513.

18-AUG-1999; 99US-0149485.

(GENE-) GENESIS RES & DEV CORP LTD.

(FLET-) FLETCHER CHALLENGE FORESTS LTD.

WOOD M, McGrath A, Shenk MA, Glenn M;

WPI; 2000-579369/54.

Claim 8; Page 422; 747pp; English.

The present invention relates to novel plant transcription factors from  
Eucalyptus grandis or Pinus radiata. The present sequence is one such  
transcription factor. The transcription factor may be used to produce a  
plant having modified gene expression such as a woody plant e.g. a  
eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or  
to modify the activity of a polypeptide in a plant. The transcription  
factors of the present invention are members from the following families  
of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic  
helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain  
zipper, LIM domain, AP2 and EREBS, zinc finger domains of type 2  
Cys2His2, CCAAT box elements and MYB.

Sequence 48 AA;

Query Match 55.6%; Score 5; DB 21; Length 48;

Best Local Similarity 100.0%; Pred. No. 79;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
|||||  
Db 24 stpva 28

## RESULT: 25

AAB44975  
ID AAB44975 standard; Protein: 57 AA.

XX AAB44975;

DT 12-FEB-2001 (first entry)

DE Human secreted protein encoded by gene 7 homologue.

XX Secreted protein: human; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KW neurotropic; neuroprotective; antibacterial; virocidic; fungicide;  
KW ophthalmological; gene therapy; treatment; autoimmune disease; infection;  
KW cardiovascular disorder; nervous system disorder; ocular disorder;  
KW wound healing; epithelial cell proliferation; skin aging; mental state;  
KW transplantation; metabolism modulation.

XX Homo sapiens.

XX WO200055200-A1.

XX 21-SEP-2000.

XX 09-MAR-2000; 2000WO-US06042.

XX 12-MAR-1999; 99US-0124143.

XX 03-DEC-1999; 99US-0168663.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-656008/63.

PT Isolated human secretory proteins, nucleic acids encoding them and  
PT antibodies directed against them, useful for diagnosing and treating  
PT disorders related to the proteins such as cancer, Alzheimer's disease  
PT and Parkinsons -

PS Disclosure: Page 405-406; 453pp; English.

XX This invention describes a novel isolated polypeptide (I) and its  
CC encoding nucleic acid molecule (II) which have immunosuppressive,  
CC antiarthritic, antirheumatic, antiproliferative, cytostatic, cardiant,  
CC vasotropic, cerebroprotective, neurotropic, neuroprotective, antibacterial,  
CC virocidic, fungicide and ophthalmological activity and which can be used  
CC for gene therapy. (I) and (II) are used to prevent, treat or ameliorate  
CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,  
CC dogs, chickens or sheep. (I) and (II) are also used in diagnosing a  
CC pathological condition or susceptibility to a pathological condition. The  
CC antibodies to (II) can also be used in alleviating symptoms associated  
CC with the disorders and in diagnostic immunoassays e.g. radioimmunoassays  
CC or enzyme linked immunosorbent assays (ELISA). Disorders which are  
CC diagnosed or treated include autoimmune diseases e.g. rheumatoid  
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or  
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular  
CC disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders  
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
CC fungi and ocular disorders e.g. corneal infection. The polypeptides can  
CC also be used to aid wound healing and epithelial cell proliferation, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides are used to  
CC modulate mammalian metabolism, to change mammal's mental state or

CC physical state by influencing biorhythms circadian rhythms, depression  
CC tendency for violence tolerance for pain, reproductive capabilities,  
CC hormonal or endocrine levels, appetite, libido, memory, stress or other  
CC cognitive qualities, as a food additive or preservative, such as to  
CC increase or decrease storage capabilities, fat content, lipid, protein,  
CC carbohydrates, vitamins, minerals, cofactors or other nutritional  
CC components.

SQ Sequence 57 AA;

Query Match 55.6%; Score 5; DB 21; Length 57;  
Best Local Similarity 100.0%; Pred. No. 93;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
|||||  
Db 29 vstpv 33

## RESULT: 26

AAB44976  
ID AAB44976 standard; Protein: 59 AA.

XX AAB44976;

DT 12-FEB-2001 (first entry)

DE Human secreted protein encoded by gene 7 homologue.

XX Secreted protein: human; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KW neurotropic; neuroprotective; antibacterial; virocidic; fungicide;  
KW ophthalmological; gene therapy; treatment; autoimmune disease; infection;  
KW cardiovascular disorder; nervous system disorder; ocular disorder;  
KW wound healing; epithelial cell proliferation; skin aging; mental state;  
KW transplantation; metabolism modulation.

XX Homo sapiens.

XX WO200055200-A1.

XX 21-SEP-2000.

XX 09-MAR-2000; 2000WO-US06042.

XX 12-MAR-1999; 99US-0124143.

XX 03-DEC-1999; 99US-0168663.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-656008/63.

PT Isolated human secretory proteins, nucleic acids encoding them and  
PT antibodies directed against them, useful for diagnosing and treating  
PT disorders related to the proteins such as cancer, Alzheimer's disease  
PT and Parkinsons -

PS Disclosure: Page 406; 453pp; English.

XX This invention describes a novel isolated polypeptide (I) and its  
CC encoding nucleic acid molecule (II) which have immunosuppressive,  
CC antiarthritic, antirheumatic, antiproliferative, cytostatic, cardiant,  
CC vasotropic, cerebroprotective, neurotropic, neuroprotective, antibacterial,  
CC virocidic, fungicide and ophthalmological activity and which can be used  
CC for gene therapy. (I) and (II) are used to prevent, treat or ameliorate  
CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,  
CC dogs, chickens or sheep. (I) and (II) are also used in diagnosing a  
CC pathological condition or susceptibility to a pathological condition. The  
CC antibodies to (II) can also be used in alleviating symptoms associated  
CC with the disorders and in diagnostic immunoassays e.g. radioimmunoassays

or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides are used to modulate mammalian metabolism, to change mammal's mental state or physical state by influencing biorhythms circadian rhythms, depression tendency for violence tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, stress or other cognitive qualities, as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrates, vitamins, minerals, cofactors or other nutritional components.

Sequence 59 AA;

Query Match 55.6%; Score 5; DB 21; Length 59;  
Best Local Similarity 100.0%; Pred. No. 96;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSPVPV 5

|||||

31 vs1pv 35

## RESULT 27

AAB39069 standard; Protein; 72 AA.

ID AAB39069;

AAB39069;

02-FEB-2001 (first entry)

Human secreted protein BLAST search protein SEQ ID NO: 159.

Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor; vulnereary; anticonvulsant; antibacterial; antifungal; antiparasitic; KM candidant; gene therapy; cancer; immune disorder; cardiovascular disorder; KM neurological disease; infection; human; secreted protein.

Homo sapiens.

MO200056880-A1.

28-SEP-2000.

16-MAR-2000; 2000MO-US06781.

19-MAR-1999; 99US-0125363.

08-DEC-1999; 99US-0169617.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Komatsoulis G;

WPI; 2000-602220/57.

Nucleic acid molecules encoding human secreted proteins, used in preventing, treating or ameliorating disorders such as Parkinson's and Alzheimer's diseases, cancers and infections -

Disclosure: Pages 409-410; 422pp: English.

The invention relates to the isolation of genes AAC59679-C59728 encoding the human secreted proteins AAB38971-B39020. This sequence represents a

peptide fragment homologous to the protein encoded by the gene isolated in the present invention. The sequence is a search result from a BLASTX CC homology search. The genes and proteins are useful for preventing, CC ameliorating or treating medical conditions, e.g. by protein or gene CC therapy. The genes are isolated from a range of human tissues disclosed CC in the specification. The nucleic acids, proteins, antibodies and CC (ant)agonists are useful in the diagnosis, treatment and prevention of: CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, CC lung, or urogenital; (b) immune disorders e.g. Addison's disease, CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid CC arthritis and ulcerative colitis; (c) cardiovascular disorders such CC as myocardial ischaemias; (d) wound healing; (e) neurological diseases CC e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such CC as viral, bacterial, fungal and parasitic infections.

Sequence 72 AA;

Query Match 55.6%; Score 5; DB 21; Length 72;  
Best Local Similarity 100.0%; Pred. No. 12e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSPVPV 5

|||||

34 vs1pv 38

## RESULT 28

AAV27569 standard; Protein; 87 AA.

ID AAV27569;

AAV27569;

30-JUL-1999 (first entry)

Human secreted protein encoded by gene No. 3.

Human; secreted protein; fusion protein; gene therapy; protein therapy; KM diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia; KM developmental abnormality; foetal deficiency; blood; allergy; renal; KM immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma; KM inflammation; ischaemic shock; Alzheimer's disease; thymus; KM cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus; KM osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion; KM endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

Homo sapiens.

WC924836-A1.

20-MAY-1999.

04-NOV-1998; 98MO-US23435.

17-NOV-1997; 97US-0066100.

07-NOV-1997; 97US-0064900.

07-NOV-1997; 97US-0064908.

07-NOV-1997; 97US-0064911.

07-NOV-1997; 97US-0064912.

07-NOV-1997; 97US-0064983.

07-NOV-1997; 97US-0064984.

07-NOV-1997; 97US-0064985.

07-NOV-1997; 97US-0064988.

17-NOV-1997; 97US-0066090.

17-NOV-1997; 97US-0066094.

17-NOV-1997; 97US-0066095.

17-NOV-1997; 97US-0066089.

(HUMA-) HUMAN GENOME SCI INC.



PI Carter KC, Ebner R, Endress GA, Feng P, Janat F;  
PI Kyaw H, Lafleur DM, Moore PA, Ni J, Olsen HS, Rosen CA;  
PI Ruben SM, Shi Y, Soppet DR, Wei Y;  
XX WPI: 1999-337740/28.  
DR N-PSDB; AAX84935.  
XX  
XX  
PT New human secreted proteins and coding sequences useful for treating  
PT disorders of the immune system and hyperproliferative disorders  
XX  
XX  
PS Claim 11: Page 350; 507pp; English.  
XX  
XX This sequence represents a secreted human protein encoded by the gene  
CC clone detailed in the descriptor line. The gene can be used to generate  
CC fusion proteins by linking to the gene to a human immunoglobulin Fc  
CC portion (e.g. AAX84924) for increasing the stability of the fused  
CC protein as compared to the human protein only.  
CC The invention relates to 125 novel genes and their fragments (nucleic  
CC acid sequences: AAX84933-85057; amino acid sequences AAY27567-Y27933)  
CC which are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Also, pathological  
CC conditions can be diagnosed by determining the amount of the new  
CC polypeptides in a sample or by determining the presence of mutations in  
CC the new polynucleotides. Specific uses are described for each of the 125  
CC polynucleotides, based on which tissues they are most highly expressed in  
CC (see AAX84933 for described uses).  
XX  
XX  
SQ Sequence 87 AA:  
  
Query Match 55.6%; Score 5; DB 20; Length 87;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 4 PVAPT 8  
|||||  
Db 60 pvapt 64  
  
RESULT 29  
ID AAB41518  
ID AAB41518 standard; Protein: 89 AA.  
XX  
XX AAB41518:  
AC  
XX  
XX .08-FEB-2001 (first entry)  
DT  
XX  
XX Human ORFX ORF1282 polypeptide sequence SEQ ID NO:2564.  
DE  
XX  
XX Human: open reading frame; ORFX: detection; cytosolic; hepatotropic;  
KW vulnery; antipsoriatic; antiparkinsonian; noctropic; neuroprotective;  
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
KW hypotensive; dermatological; immunosuppressive; antineoplastic;  
KW antiviral; antibacterial; antifungal; antineumatic; antithyroid;  
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
KW cholesterol ester storage; systemic lupus erythematosus; infection;  
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
KW bone damage; cartilage damage; antineoplastic disease; coagulation;  
KW thrombosis; contraceptive.  
XX  
XX Homo sapiens.  
OS  
XX  
XX MO200058473-A2.  
PN  
XX  
XX 05-OCT-2000.  
PD  
XX  
XX 31-MAR-2000; 2000MO-US08621.  
PF  
XX  
XX 31-MAR-1999; 99US-0127607.  
PR

PR 02-APR-1999; 99US-0127636.  
PR 05-APR-1999; 99US-0127728.  
PR 30-MAR-2000; 2000US-0540763.  
XX  
XX  
XX (CURA-) CURAGEN CORP.  
PA  
XX  
XX  
PI Shimkets RA, Leach M;  
XX  
XX  
XX WPI: 2000-602362/57.  
DR  
XX  
XX N-PSDB; AAC75727.  
PT  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease -  
XX  
XX  
PS Claim 11: Page 1830; 5507pp; English.  
XX  
XX  
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytosolic; hepatotropic; vulnery;  
CC antipsoriatic; antiparkinsonian; noctropic; neuroprotective;  
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
CC antineoplastic; antibacterial; antiviral; antifungal; antineumatic;  
CC antithyroid; and antianaemic. The sequences can be used for determining  
CC the presence of or predisposition to, or preventing or treating  
CC pathological conditions associated with an ORFX-associated disorder. The  
CC nucleic acids can be used to express ORFX proteins in gene therapy  
CC vectors. The proteins and nucleic acids may be used to treat cancers,  
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
CC nocturnal haemoglobinuria, antineoplastic disease; to enhance  
CC coagulation; to inhibit thrombosis; and as a contraceptive.  
XX  
XX  
SQ Sequence 89 AA:  
  
Query Match 55.6%; Score 5; DB 21; Length 89;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 VSTPV 5  
|||||  
Db 19 vstpv 23  
  
RESULT 30  
ID AAR47117  
ID AAR47117 standard; peptide: 94 AA.  
XX  
XX AAR47117:  
AC  
XX  
XX 30-JUN-1994 (first entry)  
DT  
XX  
XX Complete sequence of human prostatic inhibin peptide.  
DE  
XX  
XX Prostetic inhibin peptide; cancer; adenocarcinoma; breast cancer;  
KW hyperplasia; PIF; follicle stimulating hormone; FSH; tumour;  
KW inhibition; prostate cancer; gastrointestinal tract.  
XX  
XX Homo sapiens.  
OS  
XX  
XX MO9325224-A.  
PN  
XX  
XX 23-DEC-1993.  
PD  
XX  
XX 16-JUN-1993; 93MO-CA00252.  
PF  
XX  
XX

PR 16-JUN-1992; 92US-0899535.  
XX  
PA (VETR-) VETROGEN CORP.  
XX  
PI Garde S, Panchal CJ, Sheth AR;  
XX WPI: 1994-007191/01.  
XX  
PT Prepn. contg. prostatic inhibin peptide or analogues - used for  
PT treating benign prostatic hyperplasia, adenocarcinoma or diseases  
PT with high FSH levels  
PS Disclosure: Figure 1; 48pp; English.  
XX  
XX Peptides or analogues of prostatic inhibin peptide (PIP) are used  
CC in a pharmaceutical preparation for treating patients with benign  
CC prostatic hyperplasia or adenocarcinoma, particularly adenocarcinoma  
CC of the prostate or gastrointestinal tract or breast cancer. The  
CC preparation can also be used to treat patients with diseases  
CC characterised by elevated levels of follicle stimulating hormone  
CC (FSH). The PIP and its analogues inhibit tumour growth and decrease  
CC FSH levels without adversely affecting testosterone levels.  
XX  
SQ Sequence 94 AA:  
  
Query Match 55.6%; Score 5; DB 15; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
  
OY 1 VSTPV 5  
| | | | |  
DB 53 vslpv 57  
  
RESULT: 31  
AAG54394  
ID AAG54394 standard; Protein; 98 AA.  
XX  
AC AAG54394;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 69351.  
XX  
XX  
KM Protein identification: signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123588.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.

```

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

```

```

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 55.6%; Score 5; DB 21; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 2 SRPVA 6
    |||||
DB 69 srpva 73

RESULT 32
AAV82898
ID AAV82898 standard; Protein; 101 AA.
XX
AC AAV82898;
XX
XX 24-JUL-2000 (first entry)
DT
DE CUB domain from BMP-1 protein of mouse.
XX
XX BPC-1; oncogene; oncogenic; cancer; prostate; bladder; antibody;
KM antisense; vaccine; detection; prognosis; drug screening; BMP-1;
KM mouse.
XX
OS Mus musculus.
XX
PN WO200009691-A2.
XX
PD 24-FEB-2000.
XX
PF 10-AUG-1999; 99WO-US18250.
XX
XX 10-AUG-1998; 98US-0095982.
XX
XX (UROG-) UROGENESYS INC.
PA (AFAR/) AFAR D E.
PA (HUBE/) HUBERT R. S.
PA (LEON/) LEONG K.
PA (RAIT/) RAITANO A B.
PA (SAFE/) SAFEFAN D C.
PA (JAKO/) JAKOBOVITS A.
XX
XX Afar DE, Hubert RS, Leong K, Raitano AB, Safiran DC, Jakobovits A;
PI WPI; 2000-206006/18.
XX
XX New isolated BPC-1 polypeptides, useful for developing products for the
PT diagnosis, staging, prognosis and treatment of cancers, particularly
PT prostate or bladder cancer
XX
PS Disclosure: Figure 3; 79pp; English.
XX
XX BPC-1 polypeptides and polynucleotides can be used for the detection
CC of BPC-1 polypeptides and polynucleotides in biological samples, this
CC is particularly useful for detecting cancers expressing BPC-1, e.g.

```

CC prostate cancer or bladder cancer. Antibodies directed against BPC-1  
CC or antisense polynucleotides can be used for treating such cancers.  
CC The BPC-1 polypeptides can also be used in vaccines for treating or  
CC inhibiting the development of a cancer expressing BPC-1. The  
CC polypeptides and polynucleotides can also be used for detection,  
CC prognosis, drug screening and predicting susceptibility to developing  
CC cancer. In normal human tissues BPC-1 is only expressed in certain  
CC tissues of the brain, however, it is expressed at high levels in  
CC prostate cancer cells and bladder cancer cells. The BPC-1 polypeptide  
CC comprises a CUB domain which is expressed in prostate and bladder  
CC carcinoma cells and which shows sequence similarity with CUB domains  
CC from other known proteins. This is a CUB domain from the BMP-1  
CC protein of a mouse. It shows 36.3% identity with the  
CC BPC-1 CUB domain 3 in 102 residues overlap.  
XX  
XX

SO Sequence 101 AA;

Query Match 55.6%; Score 5; DB 21; Length 101;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 VAPRQ 9  
| | | | |  
Db 23 vaprq 27

RESULT 33

AAB72875  
ID AAB72875 standard; Protein: 102 AA.

XX AAB72875;

DT 09-MAY-2001 (first entry)

XX Human p53 A76T/V122A mutated peptide.

XX Human; p53 mutation; mutant; muteln; gene therapy; supertransactivating;  
XX cancer.

XX Homo sapiens.

OS Synthetic.

XX WO200109325-A2.

XX 08-FEB-2001.

XX 28-JUL-2000; 2000WO-US20538.

XX 30-JUL-1999; 99US-0146634.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Resnick MA, Inga A;

XX WPI; 2001-123321/13.

XX New isolated mutated human p53 polypeptides for inducing toxicity in a  
PT cell, treating cancer and identifying compounds that mimic toxic or  
PT supertransactivating mutations -  
XX

PS Claim 33; Page -: 144pp; English.

CC The present invention provides a number of peptides derived from the  
CC human p53 protein and containing specific mutations. These are useful for  
CC isolating supertransactivating and toxic p53 mutants, and for inhibiting  
CC cell growth and treating cancer. The present sequence is a peptide  
CC derived from the wild-type human p53 protein shown in AAB72878.  
CC Note: The present sequence is not shown in the specification but is  
CC derived from that given in Genbank acc. no. X02469 (see AAB72878).  
XX

SO Sequence 102 AA;

Query Match 55.6%; Score 5; DB 22; Length 102;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
| | | | |  
Db 23 pvapt 27

RESULT 34

AAM56088  
ID AAM56088 standard; Protein: 104 AA.

XX AAM56088;

DT 17-AUG-1998 (first entry)

XX Murine monocytic chemoattractant protein 5.

XX Murine; mouse; monocytic chemoattractant protein; MCP-4; MCP-5; chemokine;  
XX immune response; cancer; AIDS; malaria; parasitic infection.  
XX

XX Mus spretus.

XX WO9814573-A1.

XX 09-APR-1998.

XX 30-SEP-1997; 97WO-US17900.

XX 30-SEP-1996; 96US-0027128.

XX (GEHO ) GEN HOSPITAL CORP.

XX Garcia-Zepeda E, Luster AD, Sarafi M;

XX WPI; 1998-240080/21.

XX N-PSDB; AAV28592.

XX Monocyte chemoattractant proteins, MCP-4 and MCP-3 - used to develop  
PT products for treating e.g. cancers, infections, asthma, cystic  
PT fibrosis, rhinitis, atherosclerosis or inflammatory bowel disease  
XX  
XX Claim 25; Page 54; 106pp; English.

CC The present sequence represents murine monocytic chemoattractant protein 5  
CC (MCP-5). The MCP-4 and MCP-5 have activity in stimulating chemotactic  
CC activity. The proteins can be used for treating cancers, e.g. a  
CC lymphoma (e.g. Hodgkin's lymphoma), a plasmacytoma, a melanoma, a  
CC sarcoma, a tumour within the lung or gastrointestinal tract, or  
CC infectious disease such as AIDS or malaria. Antagonists to the proteins  
CC can be used for treating e.g. asthma, chronic obstructive pulmonary  
CC disease, cystic fibrosis, sinusitis, rhinitis, atherosclerosis,  
CC glomerulonephritis, multiple sclerosis, inflammatory bowel disease,  
CC arthritis or adult respiratory distress syndrome. Infections such as  
CC parasitic infections may also be treated with a molecule that inhibits  
CC MCP-4 or MCP-5 expression.  
XX

SO Sequence 104 AA;

Query Match 55.6%; Score 5; DB 19; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
| | | | |  
Db 27 vstpv 31

RESULT 35

AAM57322

ID AAW57322 standard; Protein; 104 AA.  
XX  
XX AAW57322;  
AC  
DT 11-AUG-1998 (first entry)  
XX  
XX  
DE Mouse monocytic chemotactic protein 5.  
XX  
KM Mouse; murine; monocytic chemotactic protein 5; MCP-5; m MCP-5; diagnosis;  
KM Immune system disorder; Infection; Inflammation; allergy; tumour;  
KM cardiovascular disease.  
XX  
OS Mus sp.  
XX  
XX W09812324-A1.  
PN  
XX  
XX 26-MAR-1998.  
PD  
XX  
XX 11-SEP-1997; 97W0-US16105.  
PF  
XX  
XX 08-NOV-1996; 96US-0744419.  
PR  
XX 18-SEP-1996; 96US-0026882.  
XX  
XX (BLOO-) CENT BLOOD RES INC.  
PA  
XX  
XX Gonzalo J, Gutierrez-Ramos J;  
PI  
XX  
XX WPI; 1998-217264/19.  
DR  
XX N-PSDB; AAV30789, AAV30790.  
DR  
XX  
XX Monocyte chemotactic protein-5 - used to develop products for  
PT treating e.g. immune system disorders, infections, inflammation,  
PT allergy, cardiovascular disease or tumours  
XX  
XX  
XX Claim 28; Page 79; 101pp; English.  
XX  
XX  
XX The present sequence represents monocytic chemotactic protein 5 (MCP-5).  
CC The MCP-5 protein stimulates chemotaxis of eosinophils, monocytes and  
CC lymphocytes, but not neutrophils, and so is likely to be involved with  
CC eosinophils, monocyte- and/or lymphocyte-mediated inflammations.  
CC Products of the present invention can be used for treating e.g.  
CC bacterial, fungal or parasitic infections of for tumour cell killing,  
CC or for promoting wound healing. The products can also be used for  
CC limiting an unwanted inflammatory response or an allergic response,  
CC e.g. in inflammatory bowel disease, glomerular inflammation, lupus,  
CC membranous nephropathy, glomerulo-sclerosis, chronic hepatic  
CC inflammation, fibrotic lung disease, idiopathic pulmonary disease,  
CC adult respiratory distress syndrome, sarcoidosis, pleural effusions  
CC which occur secondary to various diseases, respiratory allergies,  
CC asthma, atherosclerosis, cardiovascular disease, arthritis,  
CC endometriosis, gingivitis inflammation, inflammatory skin conditions,  
CC delayed-type hypersensitivity responses, or allergic inflammation.  
CC The products can also be used for detection, diagnosis and drug  
CC screening.  
XX  
XX Sequence 104 AA:  
SQ

Query Match 55.6%; Score 5; DB 19; Length 104;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSRPV 5  
Db 27 vscpv 31

RESULT 36  
AAG54393  
ID AAG54393 standard; Protein; 112 AA.  
XX  
XX AAG54393;  
AC  
XX

DT 18-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 69350.  
DE  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
XX EP1033405-A2.  
PN  
XX  
XX 06-SEP-2000.  
PD  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
PF  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR  
XX 05-MAR-1999; 99US-0123180.  
PR  
XX 09-MAR-1999; 99US-0123548.  
PR  
XX 23-MAR-1999; 99US-0125788.  
PR  
XX 25-MAR-1999; 99US-0126264.  
PR  
XX 29-MAR-1999; 99US-0126785.  
PR  
XX 01-APR-1999; 99US-0127462.  
PR  
XX 06-APR-1999; 99US-0128234.  
PR  
XX 08-APR-1999; 99US-0128714.  
PR  
XX 16-APR-1999; 99US-0128845.  
PR  
XX 19-APR-1999; 99US-0130077.  
PR  
XX 21-APR-1999; 99US-0130449.  
PR  
XX 23-APR-1999; 99US-0130510.  
PR  
XX 23-APR-1999; 99US-0130891.  
PR  
XX 28-APR-1999; 99US-0131449.  
PR  
XX 30-APR-1999; 99US-0132048.  
PR  
XX 30-APR-1999; 99US-0132407.  
PR  
XX 04-MAY-1999; 99US-0132484.  
PR  
XX 05-MAY-1999; 99US-0132485.  
PR  
XX 06-MAY-1999; 99US-0132486.  
PR  
XX 06-MAY-1999; 99US-0132487.  
PR  
XX 07-MAY-1999; 99US-0132863.  
PR  
XX 11-MAY-1999; 99US-0134256.  
PR  
XX 14-MAY-1999; 99US-0134218.  
PR  
XX 14-MAY-1999; 99US-0134219.  
PR  
XX 14-MAY-1999; 99US-0134221.  
PR  
XX 14-MAY-1999; 99US-0134370.  
PR  
XX 18-MAY-1999; 99US-0134768.  
PR  
XX 19-MAY-1999; 99US-0134941.  
PR  
XX 20-MAY-1999; 99US-0135124.  
PR  
XX 21-MAY-1999; 99US-0135353.  
PR  
XX 24-MAY-1999; 99US-0135629.  
PR  
XX 25-MAY-1999; 99US-0136021.  
PR  
XX 27-MAY-1999; 99US-0136392.  
PR  
XX 28-MAY-1999; 99US-0136782.  
PR  
XX 01-JUN-1999; 99US-0137222.  
PR  
XX 03-JUN-1999; 99US-0137528.  
PR  
XX 04-JUN-1999; 99US-0137502.  
PR  
XX 07-JUN-1999; 99US-0137724.  
PR  
XX 08-JUN-1999; 99US-0138094.  
PR  
XX 10-JUN-1999; 99US-0138540.  
PR  
XX 10-JUN-1999; 99US-0138847.  
PR  
XX 14-JUN-1999; 99US-0139119.  
PR  
XX 16-JUN-1999; 99US-0139452.  
PR  
XX 16-JUN-1999; 99US-0139453.  
PR  
XX 17-JUN-1999; 99US-0139492.  
PR  
XX 18-JUN-1999; 99US-0139454.  
PR  
XX 18-JUN-1999; 99US-0139455.  
PR  
XX 18-JUN-1999; 99US-0139456.  
PR  
XX 18-JUN-1999; 99US-0139457.  
PR  
XX 18-JUN-1999; 99US-0139458.  
PR  
XX 18-JUN-1999; 99US-0139459.  
PR  
XX 18-JUN-1999; 99US-0139460.  
PR  
XX 18-JUN-1999; 99US-0139461.  
PR  
XX 18-JUN-1999; 99US-0139462.  
PR  
XX 18-JUN-1999; 99US-0139463.  
PR  
XX 18-JUN-1999; 99US-0139750.

```
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142350.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148368.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153707.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match          55.6%; Score 5; DB 21; Length 112;
Best Local Similarity 100.0%; Pred: No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 STPVA 6
      |||||
Db      83 stpva 87

RESULT 37
ID      AY11113 standard; Protein: 114 AA.
XX
AC      AY11113;
XX
XX      08-JUN-1999 (first entry)
DT
DT
XX      H. pylori ORF hp6e10363_30517031_f3_3 secreted protein.
DE
XX      Vaccine; probe; diagnostic; ORF; cell envelope protein;
KW      secreted protein; cytoplasmic protein; cellular protein.
XX
```

```

OS Helicobacter pylori.
XX
XX MO9824475-A1.
XX
XX 11-JUN-1998.
XX
XX 05-DEC-1997; 97WO-US22104.
XX
XX 14-JUL-1997; 97US-0891928.
XX 05-DEC-1996; 96US-0759625.
XX 25-MAR-1997; 97US-0823745.
XX
XX (ASTR ) ASTRA AB.
XX
XX Alm RA, Castriotta LM, Dolg PC, Kabok Z, Smith D;
XX WPI: 1998-333051/29.
XX N-PSDB; AAX30642.
XX
XX New isolated Helicobacter pylori nucleic acids - used to develop
XX products for the diagnosis, prevention and treatment of infection by
XX H. pylori and other Helicobacter species
XX
XX Claims 37, 41: Page 278-279; 339pp: English.
XX
XX Recombinant or substantially pure preparations of H. pylori polypeptides
XX are disclosed, together with the nucleic acids encoding them. In all,
XX 97 ORFs are shown. The proteins are variously cell envelope proteins,
XX cytoplasmic proteins, secreted proteins or other cellular proteins.
XX Vaccines containing the nucleic acids or proteins are claimed, as are
XX probes containing at least 8 nucleotides from the nucleic acid
XX sequences. The vaccines are useful for treating or reducing the risk of
XX H. pylori infections, and the probes can be used diagnostically for
XX detecting the presence of Helicobacter in a sample. The products are
XX also of use in screening for compounds having the ability to interfere
XX with the H. pylori life cycle or to inhibit H. pylori infection.
XX
XX Sequence 114 AA:
SQ

```

Query Match 55.6%; Score 5; DB 19; Length 114;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 2 SRPVA 6
   11111
DB 99 srpva 103

```

RESULT 38  
AAG03726  
ID AAG03726 standard; Protein; 114 AA.  
XX  
XX AAG03726;  
XX  
XX 06-OCT-2000 (first entry)  
XX  
XX Human secreted protein, SEQ ID NO: 7807.  
XX  
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
XX gene therapy; chromosome mapping.  
XX  
XX Homo sapiens.  
XX  
XX EP1033401-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 21-FEB-2000; 2000EP-0200610.  
XX  
XX 26-FEB-1999; 99US-0122487.  
XX  
XX (GEST ) GENSET.  
PA

```

XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI: 2000-500381/45.
XX N-PSDB; AAC03732.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 13; SEQ ID 7807; 71pp + CD-ROM; English.
XX
XX The present sequence is a polypeptide encoded by one of a large number
XX of 5' ESTs derived from total human RNAs encoding secreted proteins. The 5' ESTs
XX were prepared from total human RNAs or polyA+ RNAs derived from 30
XX different tissues. EST sequences usually correspond mainly to the 3'
XX untranslated region (UTR) of the mRNA because they are often obtained
XX from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
XX isolating cDNA sequences derived from the 5' ends of mRNAs and even in
XX those cases where longer cDNA sequences have been obtained, the full 5'
XX UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
XX ends and can therefore be used to obtain full length cDNAs and genomic
XX DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
XX chromosome mapping procedures. They are used to obtain upstream
XX regulatory sequences and to design expression and secretion vectors.
XX
XX Sequence 114 AA:
SQ

```

Query Match 55.6%; Score 5; DB 21; Length 114;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 VSRPV 5
   11111
DB 73 vsrpv 77

```

RESULT 39  
AAG54455  
ID AAG54455 standard; Protein; 121 AA.  
XX  
XX AAG54455;  
XX  
XX 18-OCT-2000 (first entry)  
XX  
XX Zea mays protein fragment SEQ ID NO: 69435.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence; corn.  
XX  
XX Zea mays subsp. mays.  
XX  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX

PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.



PR 14-OCT-1999; 990S-0159638.  
PR 18-OCT-1999; 990S-0159584.  
PR 21-OCT-1999; 990S-0160741.  
PR 21-OCT-1999; 990S-0160767.  
PR 21-OCT-1999; 990S-0160768.  
PR 21-OCT-1999; 990S-0160770.  
PR 21-OCT-1999; 990S-0160814.  
PR 21-OCT-1999; 990S-0160815.  
PR 22-OCT-1999; 990S-0160980.  
PR 22-OCT-1999; 990S-0160981.  
PR 22-OCT-1999; 990S-0160989.  
PR 25-OCT-1999; 990S-0161404.  
PR 25-OCT-1999; 990S-0161405.  
PR 25-OCT-1999; 990S-0161406.  
PR 26-OCT-1999; 990S-0161359.  
PR 26-OCT-1999; 990S-0161360.  
PR 26-OCT-1999; 990S-0161361.  
PR 28-OCT-1999; 990S-0161920.  
PR 28-OCT-1999; 990S-0161992.  
PR 28-OCT-1999; 990S-0161993.  
PR 29-OCT-1999; 990S-0162142.

Query Match 55.6%; Score 5; DB 21; Length 121;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
Db 64 pvapt 68

RESULT 40  
AAB40264  
ID AAB40264 standard; Protein: 125 AA.

AC AAB40264;  
XX  
DT 08-FEB-2001 (first entry)  
XX

DE Human ORFX ORF28 polypeptide sequence SEQ ID NO:56.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
XX vulnery; antiporiatic; antiparkinsonian; noctropic; neuroprotective;  
XX anticonvulsant; osteopathic; antirheumatic; immunosuppressant; cardiant;  
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
XX hypotensive; dermatological; immunosuppressive; antinflammatory;  
XX antiviral; antibacterial; antifungal; antineumatic; antithyroid;  
XX antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;  
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
XX cholesterol ester storage; systemic lupus erythematosus; infection;  
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
XX bone damage; cartilage damage; antinflammatory disease; coagulation;  
XX thrombosis; contraceptive.

OS Homo sapiens.

XX WO200058473-A2.

XX PD 05-OCT-2000.

XX PF 31-MAR-2000; 2000MO-US08621.

XX PR 31-MAR-1999; 990S-0127607.

XX PR 02-APR-1999; 990S-0127636.

XX PR 05-APR-1999; 990S-0127728.

XX PR 30-MAR-2000; 2000US-0540763.

XX PA (CURA-) CURAGEN CORP.  
XX Shinkets RA, Leach M;  
XX

DR WPI, 2000-602362/57.  
DR N-PSDB; AAC74473.  
XX  
PT Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease -  
XX  
PS Claim 11; Page 449; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;  
CC antiporiatic; antiparkinsonian; noctropic; neuroprotective;  
CC osteopathic; anticonvulsant; antirheumatic; immunosuppressant;  
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
CC antinflammatory; antibacterial; antiviral; antifungal; antineumatic;  
CC antithyroid; and antianaemic. The sequences can be used for determining  
CC the presence of or predisposition to, or preventing or treating  
CC pathological conditions associated with an ORFX-associated disorder. The  
CC nucleic acids can be used to express ORFX proteins in gene therapy  
CC vectors. The proteins and nucleic acids may be used to treat cancers,  
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
CC nocturnal haemoglobinuria, antinflammatory disease; to enhance  
CC coagulation; to inhibit thrombosis; and as a contraceptive.

SO Sequence 125 AA;

Query Match 55.6%; Score 5; DB 21; Length 125;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
Db 14 tpvap 18

RESULT 41  
AAG54454  
ID AAG54454 standard; Protein: 125 AA.

XX AAG54454;

XX DT 18-OCT-2000 (first entry)

XX DE Zea mays protein fragment SEQ ID NO: 69434.

XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence; corn.

XX Zea mays subsp. mays.

XX EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 990S-0121825.

XX PR 05-MAR-1999; 990S-0123180.

XX PR 09-MAR-1999; 990S-0123548.

XX PR 23-MAR-1999; 990S-0125788.

XX PR 25-MAR-1999; 990S-0126264.

XX PR 29-MAR-1999; 990S-0126785.

XX PR 01-APR-1999; 990S-0127462.

XX PR 06-APR-1999; 990S-0128234.

PR 08-APR-1999; 9905-0128714.  
PR 16-APR-1999; 9905-0129845.  
PR 19-APR-1999; 9905-0130077.  
PR 21-APR-1999; 9905-01300449.  
PR 23-APR-1999; 9905-0130510.  
PR 28-APR-1999; 9905-0130891.  
PR 30-APR-1999; 9905-0131449.  
PR 30-APR-1999; 9905-0132048.  
PR 04-MAY-1999; 9905-0132484.  
PR 05-MAY-1999; 9905-0132485.  
PR 06-MAY-1999; 9905-0132486.  
PR 06-MAY-1999; 9905-0132487.  
PR 07-MAY-1999; 9905-0132863.  
PR 11-MAY-1999; 9905-0134256.  
PR 14-MAY-1999; 9905-0134218.  
PR 14-MAY-1999; 9905-0134219.  
PR 14-MAY-1999; 9905-0134221.  
PR 14-MAY-1999; 9905-0134370.  
PR 18-MAY-1999; 9905-0134768.  
PR 19-MAY-1999; 9905-0134941.  
PR 20-MAY-1999; 9905-0135124.  
PR 21-MAY-1999; 9905-0135353.  
PR 24-MAY-1999; 9905-0135629.  
PR 25-MAY-1999; 9905-0136021.  
PR 27-MAY-1999; 9905-0136782.  
PR 28-MAY-1999; 9905-0136782.  
PR 01-JUN-1999; 9905-0137222.  
PR 03-JUN-1999; 9905-0137528.  
PR 04-JUN-1999; 9905-0137502.  
PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138547.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139452.  
PR 16-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.  
PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140635.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141842.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142977.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0143624.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.

PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144884.  
PR 21-JUL-1999; 9905-0144814.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145085.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145218.  
PR 23-JUL-1999; 9905-0145224.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 27-JUL-1999; 9905-0145919.  
PR 28-JUL-1999; 9905-0145951.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
PR 02-AUG-1999; 9905-0146389.  
PR 03-AUG-1999; 9905-0147038.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.  
PR 06-AUG-1999; 9905-0147416.  
PR 06-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147935.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148365.  
PR 13-AUG-1999; 9905-0148684.  
PR 16-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149175.  
PR 18-AUG-1999; 9905-0149426.  
PR 20-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149723.  
PR 20-AUG-1999; 9905-0149929.  
PR 23-AUG-1999; 9905-0149902.  
PR 23-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150566.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151085.  
PR 27-AUG-1999; 9905-0151086.  
PR 27-AUG-1999; 9905-0151087.  
PR 30-AUG-1999; 9905-0151080.  
PR 31-AUG-1999; 9905-0151438.  
PR 01-SEP-1999; 9905-0151930.  
PR 07-SEP-1999; 9905-0152363.  
PR 10-SEP-1999; 9905-0153070.  
PR 13-SEP-1999; 9905-0153758.  
PR 15-SEP-1999; 9905-0154018.  
PR 16-SEP-1999; 9905-0154039.  
PR 20-SEP-1999; 9905-0154779.  
PR 22-SEP-1999; 9905-0155139.  
PR 23-SEP-1999; 9905-0155486.  
PR 24-SEP-1999; 9905-0155659.  
PR 28-SEP-1999; 9905-0156458.  
PR 29-SEP-1999; 9905-0156596.  
PR 04-OCT-1999; 9905-0157117.  
PR 05-OCT-1999; 9905-0157753.  
PR 06-OCT-1999; 9905-0157865.  
PR 07-OCT-1999; 9905-0158029.  
PR 08-OCT-1999; 9905-0158232.  
PR 12-OCT-1999; 9905-0158369.  
PR 13-OCT-1999; 9905-0159293.  
PR 13-OCT-1999; 9905-0159294.

PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159337.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161922.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 55.6%; Score 5; DB 21; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
 |||||  
 Db 68 vapt 72

## RESULT 42

AAW77657  
 ID AAW77657 standard; Protein; 130 AA.

AC AAW77657;

DT 30-OCT-1998 (first entry)

DE Staphylococcus aureus protein of unknown function.

XX Staphylococcus aureus protein; immune response induction; eye infection;  
 KW antibody production; T-cell immune response; gastrointestinal infection;  
 KW respiratory infection; inhibitor; bacterial infection; cardiac infection;  
 KW central nervous system; kidney infection; urinary tract infection;  
 KW antimicrobial compound identification; broad spectrum antibiotic;  
 KW therapy.

OS Staphylococcus aureus.

XX Key Location/Qualifiers

FT Misc-difference 1.130 /note="residues designated X are unspecified, and  
 FT represented as Xaa in the specification"

PN EP841394-A2.

PD 13-MAY-1998.

XX 24-SEP-1997; 97EP-0307485.

PR 24-SEP-1996; 96US-0027032.

XX (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.

XX Black MT, Burnham MKR, Hodgson JE, Knowles DJC;  
 PI Lonetto MA, Nicholas RO, Pratt JM, Relchard RM, Rosenberg M;

PI Ward JM;  
 XX WPI: 1998-252940/23.  
 DR N-PSDB; AAV53451.  
 DR  
 XX New nucleic acid sequences from Staphylococcus aureus MCHU29 -  
 PT useful in vaccines and for treatment of bacterial infections of e.g.  
 PT respiratory tract and central nervous system  
 XX  
 PS Claim 11; Page 312; 390pp; English.

XX This sequence represents a Staphylococcus aureus protein of unknown  
 CC function, and is encoded by a DNA sequence of the invention.  
 CC The DNA sequences were isolated from Staphylococcus aureus MCHU29  
 CC (NCIMB 40771). Host cells containing the DNA sequences are used to  
 CC produce polypeptides or fragments. The proteins are used in the treatment  
 CC of disease, for inducing an immune response by administering them, to  
 CC produce antibody and/or T-cell immune response. Antagonists of the  
 CC proteins are used for the inhibition of bacterial polypeptides.  
 CC Conditions which may be treated include bacterial infections, especially  
 CC respiratory, cardiac, gastrointestinal, central nervous, eye, kidney,  
 CC urinary tract, skin, bones and joints. The proteins can also be used to  
 CC identify antimicrobial compounds which are broad spectrum antibiotics,  
 CC especially useful in the treatment of H. pylori infection.

SO Sequence 130 AA;

Query Match 55.6%; Score 5; DB 19; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 VAPTQ 9  
 |||||  
 Db 108 vaptq 112

## RESULT 43

AAB53631  
 ID AAB53631 standard; Protein; 130 AA.

AC AAB53631;

DT 09-MAR-2001 (first entry)

DE Human colon cancer antigen protein sequence SEQ ID NO:1171.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 KW identification; cytostatic; cardioactive; neuroprotective; vulnary;  
 KW immunomodulatory; muscular; gynaecological; gastrointestinal;  
 KW neoplastic; anti-infective; antibacterial; gene therapy; wound;  
 KW neural disorder; immune system disorder; muscular disorder;  
 KW reproductive disorder; gastrointestinal disorder; renal disorder;  
 KW infectious disease; cardiovascular disorder.

XX Homo sapiens.

PN WO200055351-A1.

PD 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05883.

PR 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI: 2000-587534/55.

XX N-PSDB; AAC98388.

XX Colon cancer associated gene sequences, referred to as colon cancer

PT antigens, useful for the treatment, prevention, and diagnosis of colon  
 PT disorders such as colon cancer -

PS Claim 11: Page 1753; 2104pp; English.

XX AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
 CC called human colon cancer antigens, given in AAB53234 to AAB54006. The  
 CC human colon cancer antigens can have cytosolic, cardiolitic, muscular;  
 CC neuroprotective, immunomodulatory, gynaecological and antibacterial activities,  
 CC vulnertropic, nephrotropic, antinfecitive and antibacterial activities, and  
 CC can be used in gene therapy. The colon cancer antigen polynucleotides,  
 CC proteins and antibodies to the proteins are useful for the prevention,  
 CC treatment and diagnosis of colon disorders, such as colon cancer. The  
 CC polynucleotides may be used in diagnostics and research, such as for  
 CC chromosome identification, and as hybridisation probes. The proteins  
 CC may also be used to prevent diseases such as neural disorders, immune  
 CC system disorders, muscular disorders, reproductive disorders,  
 CC gastrointestinal disorders, wounds, renal disorders, infectious  
 CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and  
 CC AAB54007 represent sequences used in the exemplification of the present  
 CC invention.

XX Sequence 130 AA;

SQ

Query Match 55.6%; Score 5; DB 21; Length 130;  
 Best Local Similarity 100.0%; Pred. No. 2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 |||||  
 Db 63 vslpv 67

RESULT 44  
 AAB56449 standard; Protein: 132 AA.

ID AAB56449  
 XX

AC AAB56449;  
 XX

DT 13-MAR-2001 (first entry)  
 XX

DE Human prostate cancer antigen protein sequence SEQ ID NO:1027.  
 XX

XX Human prostate cancer; prostate cancer antigen; detection; diagnosis;  
 KW neuroprotective; cytosolic; cardiolitic; immunomodulatory; muscular;  
 KW vulnertropic; gastrointestinal; nephrotropic; antinfecitive; gynaecological;  
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
 KW wound; infectious disease.

XX Homo sapiens.  
 OS  
 XX

PN WO200055174-A1.  
 XX

PD 21-SEP-2000.  
 XX

PF 08-MAR-2000; 2000WO-US05988.  
 XX

PR 12-MAR-1999; 99US-0124270.  
 XX

PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX

PI Rosen CA, Ruben SM;  
 XX

DR WPI: 2000-587513/55.  
 DR N-PSDB: AAF15652.  
 XX

PT Prostate cancer associated gene sequences, referred to as prostate  
 PT cancer antigens, useful for treatment, prevention, and diagnosis of  
 PT disorders such as prostate cancer -

PS Claim 11: Page 1457; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated  
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
 CC The prostate cancer antigens can have neuroprotective, cytosolic,  
 CC cardiolitic, immunomodulatory, muscular, vulnertropic, gastrointestinal,  
 CC nephrotropic, antinfecitive, gynaecological and antibacterial activities,  
 CC and can be used in gene therapy. The prostate cancer antigen  
 CC polynucleotides may be used for detection of prostate cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The prostate cancer antigens may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
 CC AAB57303 represent sequences used in the exemplification of the present  
 CC invention.

XX Sequence 132 AA;

SQ

Query Match 55.6%; Score 5; DB 21; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 |||||  
 Db 91 vslpv 95

RESULT 45  
 AAB43699 standard; Protein: 151 AA.

ID AAB43699  
 XX

AC AAB43699;  
 XX

DT 08-FEB-2001 (first entry)  
 XX

DE Human cancer associated protein sequence SEQ ID NO:1144.  
 XX

XX Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytosolic; proliferative; vulnertropic; immunomodulatory;  
 KW antidiabetic; antistomatitic; antirheumatic; antiallergic; antiviral;  
 KW antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiant;  
 KW dermatological; neuroprotective; thrombolytic; coagulant; nocotropic;  
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;  
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening.

XX Homo sapiens.  
 OS  
 XX

PN WO200055350-A1.  
 XX

PD 21-SEP-2000.  
 XX

PF 08-MAR-2000; 2000WO-US05882.  
 XX

PR 12-MAR-1999; 99US-0124270.  
 XX

PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX

PI Rosen CA, Ruben SM;  
 XX

DR WPI: 2000-587533/55.  
 DR N-PSDB: AAC77908.  
 XX

PT Novel isolated nucleic acids comprising sequences encoding peptides  
 PT useful for treating or diagnosing e.g. cancer -

PS Claim 11: Page 1760-1761; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given

CC In ABA4398 to ABA44239. The proteins can have activities based on the  
CC tissues and cells the genes are expressed in. Example of activities  
CC include: cytostatic; proliferative; vulnerary; immunomodulator;  
CC antidiabetic; antihypertensive; antirheumatic; antarthritic;  
CC antiinflammatory; antihypertensive; antidiabetic; antiviral;  
CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;  
CC neurotrophic; vasotropic; antipsoriatic and antiangiogenic. The  
CC polynucleotides and polypeptides can be used for preventing, treating or  
CC ameliorating medical conditions and diagnosing pathological conditions  
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
CC the present invention may be used to treat immune disorders by activating  
CC or inhibiting the proliferation, differentiation or mobilisation of  
CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
CC disorders, allergic reactions, graft versus host disease and organ  
CC rejection, modulate haemostatic or thrombolytic activity, modulate  
CC inflammation, cancers, cardiovascular disorders, neurological disease and  
CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
CC AAC78457 and ABA44240 represent sequences used in the exemplification of  
CC the present invention.

SO Sequence 151 AA:

Query Match 55.6%; Score 5; DB 21; Length 151;  
Best Local Similarity 100.0%; Pred. No. 2.36+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
|||||  
DB 50 tpvap 54

## RESULT 46

AAG54453  
ID AAG54453 standard; Protein; 152 AA.

AC AAG54453;

XX 18-OCT-2000 (first entry)

XX Zea mays protein fragment SEQ ID NO: 69433.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.

XX Zea mays subsp. mays.

PN EP103405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131445.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
PR 01-JUN-1999; 99US-0136782.  
PR 03-JUN-1999; 99US-0137222.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137502.  
PR 08-JUN-1999; 99US-0137724.  
PR 10-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.

[illegible]

PR	09-JUL-1999;	99DE-1032125.
PR	09-JUL-1999;	99DE-1032126.
PR	09-JUL-1999;	99DE-1032130.
PR	09-JUL-1999;	99DE-1032186.
PR	09-JUL-1999;	99DE-1032206.
PR	09-JUL-1999;	99DE-1032227.
PR	09-JUL-1999;	99DE-1032228.
PR	09-JUL-1999;	99DE-1032229.
PR	09-JUL-1999;	99DE-1032230.
PR	14-JUL-1999;	99DE-1032922.
PR	14-JUL-1999;	99DE-1032926.
PR	14-JUL-1999;	99DE-1032928.
PR	14-JUL-1999;	99DE-1033004.
PR	14-JUL-1999;	99DE-1033005.
PR	14-JUL-1999;	99DE-1033006.
PR	12-AUG-1999;	99US-0148613.
PR	27-AUG-1999;	99DE-1040764.
PR	27-AUG-1999;	99DE-1040765.
PR	27-AUG-1999;	99DE-1040766.
PR	27-AUG-1999;	99DE-1040832.
PR	31-AUG-1999;	99DE-1041378.
PR	31-AUG-1999;	99DE-1041379.
PR	31-AUG-1999;	99DE-1041380.
PR	31-AUG-1999;	99DE-1041394.
PR	31-AUG-1999;	99DE-1041396.
PR	03-SEP-1999;	99DE-1042076.
PR	03-SEP-1999;	99DE-1042077.
PR	03-SEP-1999;	99DE-1042079.
PR	03-SEP-1999;	99DE-1042086.
PR	03-SEP-1999;	99DE-1042087.
PR	03-SEP-1999;	99DE-1042088.
PR	03-SEP-1999;	99DE-1042095.
PR	03-SEP-1999;	99DE-1042124.
PR	03-SEP-1999;	99DE-1042129.
PR	09-MAR-2000;	2000US-0187970.
XX	(BADI ) BASF AG.	
PA		
XX		
PI	Pompejus M, Kroegeer B, Schroeder H, zelder O, Haberhauer G;	
XX		
DR	WPI: 2001-137957/14.	
DR	N-PSDB; AAF72092.	
XX		
PT	Nucleic acids from Corynebacterium glutamicum encoding metabolic	
PT	pathway proteins, useful for producing fine chemicals in	
PT	microorganisms, including organic acids, nonproteinogenic amino acids,	
PT	and purine and pyrimidine bases -	
XX		
PS	Claim 20; Page 1107; 1737pp; English.	
XX		
CC	AAE71753 to AAE72330 encode the Corynebacterium glutamicum metabolic	
CC	pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum	
CC	MP nucleic acids are useful for the production of fine chemicals	
CC	In microorganisms, including organic acids, nonproteinogenic amino	
CC	acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,	
CC	saturated and unsaturated fatty acids, diols, carbohydrates, aromatic	
CC	compounds, vitamins, cofactors, polyketides and enzymes.	
XX		
SO	Sequence 152 AA;	
QY	Query Match 55.6%; Score 5; DB 22; Length 152;	
	Best Local Similarity 100.0%; Pred. No. 2,4e+02;	
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
DB	115 vstpv 119	
RESULT 48		
AAE79981		
ID	AAE79981 standard; Protein: 152 AA.	

XX	AAE79981:	
XX		
XX	30-APR-2001	(first entry)
XX		
DE	Corynebacterium glutamicum	MP protein sequence SEQ ID NO:696.
XX		
KW	Corynebacterium glutamicum; metabolic pathway protein; MP protein;	
KW	fine chemical production; microorganism; organic acid; nucleoside;	
KW	nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;	
KW	lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;	
KW	carbohydrate; aromatic compound; cofactor; polyketide; enzyme.	
XX		
OS	Corynebacterium glutamicum.	
XX		
FN	MO200100843-A2.	
XX		
PD	04-JAN-2001.	
XX		
PE	23-JUN-2000;	2000MO-IB00923.
XX		
PR	25-JUN-1999;	99US-0141031.
PR	01-JUL-1999;	99DE-1030476.
PR	02-JUL-1999;	99US-0142101.
PR	08-JUL-1999;	99DE-1031415.
PR	08-JUL-1999;	99DE-1031418.
PR	08-JUL-1999;	99DE-1031419.
PR	08-JUL-1999;	99DE-1031420.
PR	08-JUL-1999;	99DE-1031424.
PR	08-JUL-1999;	99DE-1031428.
PR	08-JUL-1999;	99DE-1031434.
PR	08-JUL-1999;	99DE-1031435.
PR	08-JUL-1999;	99DE-1031443.
PR	08-JUL-1999;	99DE-1031453.
PR	08-JUL-1999;	99DE-1031457.
PR	08-JUL-1999;	99DE-1031465.
PR	08-JUL-1999;	99DE-1031478.
PR	08-JUL-1999;	99DE-1031510.
PR	08-JUL-1999;	99DE-1031541.
PR	08-JUL-1999;	99DE-1031573.
PR	08-JUL-1999;	99DE-1031592.
PR	08-JUL-1999;	99DE-1031632.
PR	08-JUL-1999;	99DE-1031634.
PR	08-JUL-1999;	99DE-1031636.
PR	09-JUL-1999;	99DE-1032125.
PR	09-JUL-1999;	99DE-1032126.
PR	09-JUL-1999;	99DE-1032130.
PR	09-JUL-1999;	99DE-1032186.
PR	09-JUL-1999;	99DE-1032206.
PR	09-JUL-1999;	99DE-1032227.
PR	09-JUL-1999;	99DE-1032228.
PR	09-JUL-1999;	99DE-1032229.
PR	09-JUL-1999;	99DE-1032230.
PR	14-JUL-1999;	99DE-1032922.
PR	14-JUL-1999;	99DE-1032926.
PR	14-JUL-1999;	99DE-1032928.
PR	14-JUL-1999;	99DE-1033004.
PR	14-JUL-1999;	99DE-1033005.
PR	14-JUL-1999;	99DE-1033006.
PR	12-AUG-1999;	99US-0148613.
PR	27-AUG-1999;	99DE-1040764.
PR	27-AUG-1999;	99DE-1040765.
PR	27-AUG-1999;	99DE-1040766.
PR	27-AUG-1999;	99DE-1040832.
PR	31-AUG-1999;	99DE-1041378.
PR	31-AUG-1999;	99DE-1041379.
PR	31-AUG-1999;	99DE-1041380.
PR	31-AUG-1999;	99DE-1041394.
PR	31-AUG-1999;	99DE-1041396.
PR	03-SEP-1999;	99DE-1042076.
PR	03-SEP-1999;	99DE-1042077.
PR	03-SEP-1999;	99DE-1042079.
PR	03-SEP-1999;	99DE-1042086.

PR 03-SEP-1999; 99DE-1042087.  
PR 03-SEP-1999; 99DE-1042088.  
PR 03-SEP-1999; 99DE-1042095.  
PR 03-SEP-1999; 99DE-1042124.  
PR 03-SEP-1999; 99DE-1042129.  
PR 09-MAR-2000; 2000US-0187970.  
XX  
XX (BADI ) BASF AG.  
PI Pompeius M, Kroegeer B, Schroeder H, Zeider O, Haberhauer G;  
XX WPI; 2001-137957/14.  
DR N-PSDB; AAF72100.  
XX  
XX Nucleic acids from *Corynebacterium glutamicum* encoding metabolic  
PT pathway proteins, useful for producing fine chemicals in  
PT microorganisms, including organic acids, nonproteinogenic amino acids,  
PT and purine and pyrimidine bases -  
XX  
XX Claim 20; Page 1123-1124; 1737pp; English.  
XX  
XX AAF71753 to AAF72330 encode the *Corynebacterium glutamicum* metabolic  
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum  
CC MP nucleic acids are useful for the production of fine chemicals  
CC in microorganisms, including organic acids, nonproteinogenic amino  
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,  
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic  
CC compounds, vitamins, cofactors, polyketides and enzymes.  
XX  
XX Sequence 152 AA;  
SQ  
  
Query Match 55.6%; Score 5; DB 22; Length 152;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 VSTRPV 5  
| | | | |  
Db 115 vstrpv 119  
  
RESULT 49  
AAB79987  
ID AAB79987 standard; Protein; 152 AA.  
XX  
XX AAB79987;  
AC  
XX  
DT 30-APR-2001 (first entry)  
XX  
XX  
DE *Corynebacterium glutamicum* MP protein sequence SEQ ID NO:708.  
XX  
XX *Corynebacterium glutamicum*; metabolic pathway protein; MP protein;  
KM fine chemical production; microorganism; organic acid; nucleoside;  
KM nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;  
KM lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;  
KM carbohydrate; aromatic compound; cofactor; polyketide; enzyme.  
XX  
XX *Corynebacterium glutamicum*.  
OS  
XX  
PN WO200100843-A2.  
XX  
PD 04-JAN-2001.  
XX  
XX  
PF 23-JUN-2000; 2000MO-IB00923.  
XX  
XX  
PR 25-JUN-1999; 99US-0141031.  
PR 01-JUL-1999; 99DE-1030476.  
PR 07-JUL-1999; 99US-0142101.  
PR 08-JUL-1999; 99DE-1031415.  
PR 08-JUL-1999; 99DE-1031418.  
PR 08-JUL-1999; 99DE-1031419.  
PR 08-JUL-1999; 99DE-1031420.  
PR 08-JUL-1999; 99DE-1031424.  
PR 08-JUL-1999; 99DE-1031424.

PR 08-JUL-1999; 99DE-1031428.  
PR 08-JUL-1999; 99DE-1031434.  
PR 08-JUL-1999; 99DE-1031435.  
PR 08-JUL-1999; 99DE-1031443.  
PR 08-JUL-1999; 99DE-1031453.  
PR 08-JUL-1999; 99DE-1031457.  
PR 08-JUL-1999; 99DE-1031478.  
PR 08-JUL-1999; 99DE-1031478.  
PR 08-JUL-1999; 99DE-1031510.  
PR 08-JUL-1999; 99DE-1031541.  
PR 08-JUL-1999; 99DE-1031573.  
PR 08-JUL-1999; 99DE-1031592.  
PR 08-JUL-1999; 99DE-1031632.  
PR 08-JUL-1999; 99DE-1031634.  
PR 08-JUL-1999; 99DE-1031636.  
PR 09-JUL-1999; 99DE-1032125.  
PR 09-JUL-1999; 99DE-1032126.  
PR 09-JUL-1999; 99DE-1032130.  
PR 09-JUL-1999; 99DE-1032186.  
PR 09-JUL-1999; 99DE-1032206.  
PR 09-JUL-1999; 99DE-1032227.  
PR 09-JUL-1999; 99DE-1032228.  
PR 09-JUL-1999; 99DE-1032229.  
PR 09-JUL-1999; 99DE-1032230.  
PR 14-JUL-1999; 99DE-1032922.  
PR 14-JUL-1999; 99DE-1032926.  
PR 14-JUL-1999; 99DE-1032928.  
PR 14-JUL-1999; 99DE-1033004.  
PR 14-JUL-1999; 99DE-1033005.  
PR 14-JUL-1999; 99DE-1033006.  
PR 12-AUG-1999; 99US-0148613.  
PR 27-AUG-1999; 99DE-1040764.  
PR 27-AUG-1999; 99DE-1040765.  
PR 27-AUG-1999; 99DE-1040766.  
PR 27-AUG-1999; 99DE-1040832.  
PR 31-AUG-1999; 99DE-1041378.  
PR 31-AUG-1999; 99DE-1041379.  
PR 31-AUG-1999; 99DE-1041380.  
PR 31-AUG-1999; 99DE-1041394.  
PR 31-AUG-1999; 99DE-1041396.  
PR 03-SEP-1999; 99DE-1042076.  
PR 03-SEP-1999; 99DE-1042077.  
PR 03-SEP-1999; 99DE-1042079.  
PR 03-SEP-1999; 99DE-1042086.  
PR 03-SEP-1999; 99DE-1042087.  
PR 03-SEP-1999; 99DE-1042088.  
PR 03-SEP-1999; 99DE-1042095.  
PR 03-SEP-1999; 99DE-1042124.  
PR 03-SEP-1999; 99DE-1042129.  
PR 09-MAR-2000; 2000US-0187970.  
XX  
XX (BADI ) BASF AG.  
PA  
XX  
XX Pompeius M, Kroegeer B, Schroeder H, Zeider O, Haberhauer G;  
PI WPI; 2001-137957/14.  
DR N-PSDB; AAF72106.  
XX  
XX Nucleic acids from *Corynebacterium glutamicum* encoding metabolic  
PT pathway proteins, useful for producing fine chemicals in  
PT microorganisms, including organic acids, nonproteinogenic amino acids,  
PT and purine and pyrimidine bases -  
XX  
XX  
PS Claim 20; Page 1141-1142; 1737pp; English.  
XX  
XX AAF71753 to AAF72330 encode the *Corynebacterium glutamicum* metabolic  
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum  
CC MP nucleic acids are useful for the production of fine chemicals  
CC in microorganisms, including organic acids, nonproteinogenic amino  
CC acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids,  
CC saturated and unsaturated fatty acids, diols, carbohydrates, aromatic  
CC compounds, vitamins, cofactors, polyketides and enzymes.  
XX



SQ Sequence 152 AA;

DB 55 vstpv 59

Query Match 55.6%; Score 5; DB 22; Length 152;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Search completed: August 15, 2001, 12:35:31

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Job time: 156 sec

OY 1 VSTPV 5

DB 115 vstpv 119

RESULT 50

AAB24690

ID AAB24690 standard; Peptide: 158 AA.

AC AAB24690;

DT 27-NOV-2000 (first entry)

DE Plant SDF encoded polypeptide sequence SEQ List 1 NO:122.

KW Plant; corn; Arabidopsis thaliana; sequence-determined DNA fragment;  
SDF; genetic mapping; identification; promoter; structural gene; UTR;  
untranslated region; expression control.

XX OS Plant.

PN WO200040695-A2.

PD 13-JUL-2000.

PE 07-JAN-2000; 2000WO-US00466.

PR 08-JAN-1999; 99US-0115293.

PA (CERE-) CERES INC.

PI Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;

PI Zheng L;

DR WPI; 2000-465970/40.

PT New corn plant and Arabidopsis thaliana sequence-determined DNA  
PT fragments, useful for expressing gene products and for controlling  
PT expression of a target gene -

PS Claim 14; Page 381-382; 673pp; English.

XX The present invention describes polynucleotides, such as complete cDNA  
CC sequences and/or sequences of genomic DNA encompassing complete genes,  
CC portions of genes, and/or intergenic regions, collectively referred to  
CC as sequence-determined DNA fragments (SDFs), from corn plants and  
CC Arabidopsis thaliana. The SDFs are promoters, structural genes, and  
CC untranslated regions (UTRs), or 3' termination sequences. They can be  
CC used for expressing a gene product and controlling expression of a  
CC target gene, either as a promoter, a structural gene, an UTR or as a  
CC 3' termination sequence. They are also useful as tools for genetic  
CC mapping, and identification of a particular individual plant or for  
CC clustering a group of plants with a common trait. AAA78433 to AAA78630  
CC and AAB24605 to AAB25099 represent the specifically claimed  
CC polynucleotide sequences and polypeptides encoded by them given in the  
CC present invention.

XX Sequence 158 AA;

Query Match 55.6%; Score 5; DB 21; Length 158;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5

IIIIII

**THIS PAGE BLANK (uspto)**

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:32:55 ; Search time 33.87 Seconds  
(without alignments)  
20.241 Million cell updates/sec

Title: US-09-372-036-26  
Perfect score: 9  
Sequence: 1 VSRPVPATQ 9

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	100.0	484	2 A41487	protein P60 precu
2	6	66.7	102	2 F83231	hypothetical prote
3	6	66.7	155	2 S35162	SRH-21 protein - p
4	6	66.7	187	2 JC4806	core protein G - p
5	6	66.7	218	2 C86847	transcription regu
6	6	66.7	288	2 F71504	hypothetical prote
7	6	66.7	293	2 F75466	hypothetical prote
8	6	66.7	447	2 T34992	probable lipoprote
9	6	66.7	453	1 F64623	amidase - Helicoba
10	6	66.7	453	1 A71891	glu-tRNA amidotran
11	6	66.7	460	2 T23087	hypothetical prote
12	6	66.7	637	2 S66953	hypothetical prote
13	6	66.7	748	2 T49633	glycan 1,4-alpha-g
14	6	66.7	748	2 H84913	probable ATP-depen
15	6	66.7	865	2 A47882	calcium-binding pr
16	6	66.7	1093	2 T50652	AP-3 complex beta3
17	6	66.7	1094	2 T50651	AP-3 complex beta3
18	6	66.7	1105	2 T18295	AP-3 adaptor compl
19	5	55.6	75	2 E75264	hypothetical prote
20	5	55.6	82	2 T29802	hypothetical prote
21	5	55.6	88	2 T34429	hypothetical prote
22	5	55.6	103	1 K4RB	Ig kappa-B4 chain
23	5	55.6	104	2 F53275	beta-microseminop
24	5	55.6	114	2 A34567	methylnalonyl-CoA
25	5	55.6	115	2 B49094	conserved hypotnet
26	5	55.6	117	2 A72293	hypothetical prote
27	5	55.6	123	2 G71189	lymphocyte antigen
28	5	55.6	126	2 I54454	hypothetical 13.4K
29	5	55.6	128	2 JN0727	

30	5	55.6	132	2 T03397	hypothetical prote
31	5	55.6	134	2 B60497	MHC class II histo
32	5	55.6	141	2 E72580	hypothetical prote
33	5	55.6	147	2 T15036	photosystem I chai
34	5	55.6	161	2 C96568	hypothetical prote
35	5	55.6	162	2 B84018	hypothetical prote
36	5	55.6	165	2 S61230	cytochrome-c blosy
37	5	55.6	170	2 S11770	lmaA protein - Lis
38	5	55.6	170	2 B83871	hypothetical prote
39	5	55.6	175	2 T50849	hypothetical prote
40	5	55.6	177	1 S57789	hypothetical prote
41	5	55.6	178	1 KRRT	kappa-casein precu
42	5	55.6	178	2 S56300	hypothetical prote
43	5	55.6	179	2 S01400	H+-transporting AT
44	5	55.6	179	2 G84789	hypothetical prote
45	5	55.6	181	2 A05080	kappa-casein precu
46	5	55.6	185	2 D96572	protein P12M16.6 I
47	5	55.6	193	2 H86404	probable lipid tra
48	5	55.6	197	2 D85875	probable minor fil
49	5	55.6	199	2 F86832	hypothetical prote
50	5	55.6	201	2 S69748	hypothetical prote
51	5	55.6	206	2 H83587	conserved hypotnet
52	5	55.6	211	2 T22122	hypothetical prote
53	5	55.6	217	2 S10212	late 33K protein -
54	5	55.6	218	2 G71260	probable ribulose-
55	5	55.6	223	2 G82478	hypothetical prote
56	5	55.6	224	2 A31459	MHC class II histo
57	5	55.6	225	2 F72642	probable shikimate
58	5	55.6	228	2 F75390	hypothetical prote
59	5	55.6	229	1 LCBO	prolactin precurs
60	5	55.6	229	1 LCSH	prolactin - goat
61	5	55.6	229	2 T83982	hypothetical prote
62	5	55.6	241	2 T17798	probable muramidase
63	5	55.6	245	2 F71474	succinate dehydrog
64	5	55.6	251	2 C84036	hypothetical prote
65	5	55.6	252	2 A86449	hypothetical prote
66	5	55.6	260	2 T18909	MHC class II histo
67	5	55.6	261	1 HLH0DB	MHC class II histo
68	5	55.6	261	1 HLH0IC	MHC class II histo
69	5	55.6	261	2 B37044	MHC class II histo
70	5	55.6	261	2 I54480	HLA DO-beta - huma
71	5	55.6	261	2 T51222	hypothetical prote
72	5	55.6	264	2 C37386	hypothetical prote
73	5	55.6	264	2 S01922	hypothetical prote
74	5	55.6	267	2 T32063	hypothetical prote
75	5	55.6	268	2 D29312	MHC class II histo
76	5	55.6	269	2 I54432	MHC class II histo
77	5	55.6	271	2 T16421	hypothetical prote
78	5	55.6	273	2 F84242	imidazoleglycerol-
79	5	55.6	276	2 S75249	esterase all0992 -
80	5	55.6	276	2 T29894	hypothetical prote
81	5	55.6	282	2 E71662	protein export pro
82	5	55.6	288	2 A81009	hypothetical prote
83	5	55.6	293	2 T35621	hypothetical prote
84	5	55.6	296	2 G84747	AT-hook DNA-bindin
85	5	55.6	298	2 T36900	probable integral
86	5	55.6	299	2 A82063	uroporphyrin-III C
87	5	55.6	302	2 A96841	hypothetical prote
88	5	55.6	303	2 S39611	phosphoprotein pho
89	5	55.6	305	2 B55346	phosphoprotein pho
90	5	55.6	307	1 S75769	hypothetical prote
91	5	55.6	307	2 T24423	hypothetical prote
92	5	55.6	311	2 H84746	hypothetical prote
93	5	55.6	314	2 S22513	ethylene-forming e
94	5	55.6	315	2 T25171	hypothetical prote
95	5	55.6	315	2 T26422	hypothetical prote
96	5	55.6	318	2 T00112	glycosyltransferas
97	5	55.6	319	2 T46145	hypothetical prote
98	5	55.6	321	2 J50772	37K glycoprotein -
99	5	55.6	323	2 A31351	probable transmemb
100	5	55.6	324	2 C83695	quinone oxidoreduc
101	5	55.6	324	2 T24465	hypothetical prote
102	5	55.6	325	1 A45470	hydroxymethylgluta

103	5	55.6	325	2	D70666	probable modb prot	176	5	55.6	471	2	JC1403	glutamate--ammonia
104	5	55.6	331	2	T28694	hypothetical prote	177	5	55.6	473	2	S75141	glutamate--ammonia
105	5	55.6	332	2	T26145	hypothetical prote	178	5	55.6	473	2	H69761	conserved hypochet
106	5	55.6	335	2	T31559	hypothetical prote	179	5	55.6	474	2	T10371	capsid-associated
107	5	55.6	335	2	T31560	hypothetical prote	180	5	55.6	475	2	T44597	conserved hypochet
108	5	55.6	335	2	T31561	hypothetical prote	181	5	55.6	475	2	D86209	protein P2265.18
109	5	55.6	335	2	H75518	probable cytochrom	182	5	55.6	475	2	A85923	hypothetical prote
110	5	55.6	336	1	B38274	y box-binding prot	183	5	55.6	476	2	B44997	merozoite surface
111	5	55.6	337	2	C69124	ribosomal protein	184	5	55.6	476	2	T25736	hypothetical prote
112	5	55.6	337	2	G75529	peptide ABC transp	185	5	55.6	477	2	G71319	hypothetical prote
113	5	55.6	342	2	T45456	probable membrane	186	5	55.6	478	2	JC4880	fibrinolytic metal
114	5	55.6	343	2	T14245	NADH dehydrogenase	187	5	55.6	478	2	A43296	atrolysin E (EC 3.
115	5	55.6	356	1	XNECHC	histidinol-phospha	188	5	55.6	478	2	A32555	major merozoite su
116	5	55.6	356	2	D85827	histidinol-phospha	189	5	55.6	480	1	A30065	trigramin precursor
117	5	55.6	358	2	T35104	probable two-compo	190	5	55.6	481	2	S43125	fibrinolytic prote
118	5	55.6	359	1	XNEBHC	histidinol-phospha	191	5	55.6	481	2	JC4342	merozoite surface
119	5	55.6	362	2	S24551	protein-tyrosine k	192	5	55.6	482	2	A44997	hypothetical prote
120	5	55.6	366	1	MNKRSD	sigma NS protein -	193	5	55.6	482	2	T32564	hypothetical prote
121	5	55.6	366	1	MNKRST	sigma NS protein -	194	5	55.6	488	1	QOBEHS	alkaline exonuclea
122	5	55.6	366	1	MNKRRT	sigma NS protein -	195	5	55.6	488	2	T44030	alkaline exonuclea
123	5	55.6	366	2	C27401	sigma NS protein -	196	5	55.6	488	2	T44215	alkaline exonuclea
124	5	55.6	366	2	S53073	hypothetical prote	197	5	55.6	488	2	F86209	protein P2265.14
125	5	55.6	367	1	A46355	site-specific DNA-	198	5	55.6	486	2	T30976	hypothetical prote
126	5	55.6	367	2	T18185	probable site-spec	199	5	55.6	488	2	G02421	hypothetical prote
127	5	55.6	368	2	S06058	NADH dehydrogenase	200	5	55.6	500	2	S77243	hypothetical prote
128	5	55.6	368	2	G84769	hypothetical prote	201	5	55.6	501	2	T27513	hypothetical prote
129	5	55.6	372	2	G01425	nucleolar 75K auto	202	5	55.6	504	2	T34995	probable integral
130	5	55.6	375	2	T51333	transcription fact	203	5	55.6	505	2	G02474	interferon regulat
131	5	55.6	376	2	S04497	surface antigen PA	204	5	55.6	505	2	B46629	mucin 6, gastric
132	5	55.6	379	2	A45443	tubulin--tyrosine	205	5	55.6	508	2	T19350	hypothetical prote
133	5	55.6	380	2	B69523	succinyl-CoA synth	206	5	55.6	509	1	S04346	steroid 17alpha-mo
134	5	55.6	380	2	D84214	NADH dehydrogenase	207	5	55.6	509	1	S36572	protein - human
135	5	55.6	383	2	S32975	gene BCRP2 protein	208	5	55.6	510	2	S53970	amidophosphoribosy
136	5	55.6	386	2	S64614	probable membrane	209	5	55.6	510	2	A45338	conexin-56 - chic
137	5	55.6	390	2	F65048	multidrug resistan	210	5	55.6	512	2	T33463	probable serine ca
138	5	55.6	390	2	F85916	multidrug resistan	211	5	55.6	513	2	D72753	probable xyulose
139	5	55.6	391	2	T51097	thyroid hormone re	212	5	55.6	514	1	P1WL8	l1 protein - human
140	5	55.6	394	2	F69888	alanine racemase h	213	5	55.6	514	1	P1WL47	l1 protein - human
141	5	55.6	394	2	T18752	hypothetical prote	214	5	55.6	516	1	P1WL5	l1 protein - human
142	5	55.6	402	2	C64895	hypothetical prote	215	5	55.6	517	2	S36496	l1 protein - human
143	5	55.6	402	2	G86117	probable virulence	216	5	55.6	517	2	S36543	l2 protein - human
144	5	55.6	402	2	G85882	probable virulence	217	5	55.6	518	2	S36472	l2 protein - human
145	5	55.6	402	2	JH0403	procollagen I C-pr	218	5	55.6	525	1	P1WL85	l1 protein - human
146	5	55.6	404	2	T35050	hypothetical prote	219	5	55.6	530	2	D83291	probable MFS trans
147	5	55.6	407	2	S66260	metalloproteinase	220	5	55.6	534	2	C69146	hypothetical prote
148	5	55.6	408	2	T37929	probable major fac	221	5	55.6	537	2	E96606	hypothetical prote
149	5	55.6	410	2	D70599	hypothetical prote	222	5	55.6	542	2	T29707	hypothetical prote
150	5	55.6	414	1	HYRSNC	atrolysin C (EC 3.	223	5	55.6	543	2	T00513	cytochrome P450 ho
151	5	55.6	414	2	S41608	atrolysin B (EC 3.	224	5	55.6	543	2	S19933	glycine-rich prote
152	5	55.6	414	2	S41609	atrolysin C (EC 3.	225	5	55.6	544	2	T40942	glycine-rich prote
153	5	55.6	415	2	T20335	hypothetical prote	226	5	55.6	544	2	T20741	hypothetical prote
154	5	55.6	416	2	G69748	conserved hypochet	227	5	55.6	546	2	S36490	carbamoyl-phosphat
155	5	55.6	418	2	F83986	transporter BH2694	228	5	55.6	547	2	B56573	probable enzyme 24
156	5	55.6	420	2	B84788	hypothetical prote	229	5	55.6	550	2	C86704	conserved hypochet
157	5	55.6	421	2	D81098	gcpe protein NMB13	230	5	55.6	553	2	C75318	hypothetical prote
158	5	55.6	421	2	H81843	conserved hypochet	231	5	55.6	556	2	D86262	hypothetical prote
159	5	55.6	422	2	G70984	probable papas pro	232	5	55.6	557	2	T47511	probable transport
160	5	55.6	423	1	XXPS2M	dihydroallopsamide s	233	5	55.6	557	2	A70480	carbamoyl-phosphat
161	5	55.6	423	2	H84257	Hrli3 transducer l	234	5	55.6	558	2	E86016	probable enzyme 24
162	5	55.6	423	2	T44258	transducer protein	235	5	55.6	559	2	C75286	hypothetical prote
163	5	55.6	423	2	S73020	hypothetical prote	236	5	55.6	570	2	F70332	proline-tRNA synth
164	5	55.6	428	2	E83365	branched-chain alp	237	5	55.6	573	2	S66710	probable membrane
165	5	55.6	432	1	A41056	brachyury homolog	238	5	55.6	579	2	S54872	penicillin-binding
166	5	55.6	442	2	D71481	probable trigger f	239	5	55.6	581	2	E75383	conserved hypochet
167	5	55.6	444	2	B84292	hypothetical prote	240	5	55.6	585	2	F85809	hypothetical prote
168	5	55.6	448	2	B56558	retinoic acid rece	241	5	55.6	587	2	F85084	probable athlia-11
169	5	55.6	451	2	G83916	hypocotate oxidase	242	5	55.6	599	2	D81367	probable pyruvate
170	5	55.6	453	2	S40943	hypocotate prote	243	5	55.6	601	1	A64222	heat shock protein
171	5	55.6	455	2	H83494	probable 2-isoprop	244	5	55.6	603	2	S06059	gene NDI Intron 4
172	5	55.6	456	2	H85574	hypocotate prote	245	5	55.6	604	2	S54032	probable amino acil
173	5	55.6	459	2	A41977	retinoic acid rece	246	5	55.6	604	2	S25203	srnr protein - Str
174	5	55.6	462	2	B86262	hypocotate prote	247	5	55.6	605	2	E69153	cadmium efflux ATP
175	5	55.6	464	2	A56558	retinoic acid rece	248	5	55.6	609	2	E72012	glutamine--fructos

249	5	55.6	609	2	S45930	probable amino aci
250	5	55.6	612	2	G71972	threonine--tRNA L
251	5	55.6	612	2	C64535	threonine--tRNA L
252	5	55.6	613	2	G82762	penicillin binding
253	5	55.6	614	2	A56614	acetate--CoA ligas
254	5	55.6	622	2	A64980	hypothetical 68.5k
255	5	55.6	634	1	W1WL51	E1 protein - human
256	5	55.6	634	2	I49642	estrogen-responsiv
257	5	55.6	637	2	S74677	hypothetical prote
258	5	55.6	643	1	DJEC36	DNA-directed DNA p
259	5	55.6	643	2	G85344	hypothetical prote
260	5	55.6	648	2	F85845	unknown protein en
261	5	55.6	648	2	T37581	probable serine-th
262	5	55.6	648	2	PC4395	myosin 3 - human (f
263	5	55.6	649	2	JN0809	drebrin E (clone g
264	5	55.6	651	2	T15624	hypothetical prote
265	5	55.6	651	2	T14763	hypothetical prote
266	5	55.6	662	2	T18233	probable transcript
267	5	55.6	670	2	T34548	hypothetical prote
268	5	55.6	671	2	A38109	autolysin - entero
269	5	55.6	682	2	T47473	receptor-like prot
270	5	55.6	682	2	C84295	UDP-sugar hydrolas
271	5	55.6	683	2	E86358	hypothetical prote
272	5	55.6	685	3	JC7570	Delta-4 protein -
273	5	55.6	686	3	JC7569	hypothetical prote
274	5	55.6	700	2	D85582	hypothetical prote
275	5	55.6	700	2	T33577	hypothetical prote
276	5	55.6	701	1	S33709	DHR39--short protei
277	5	55.6	707	2	JC2218	procollagen C-endo
278	5	55.6	711	2	A85352	cadmium-transporti
279	5	55.6	730	1	BMH01	procollagen C-endo
280	5	55.6	748	2	S54505	hypothetical prote
281	5	55.6	750	1	H70628	probable pkng prot
282	5	55.6	753	2	T24869	hypothetical prote
283	5	55.6	755	2	S42462	structural polypyo
284	5	55.6	763	2	T44735	serine/threonine-s
285	5	55.6	767	2	S63220	probable membrane
286	5	55.6	768	2	T37601	probable transcript
287	5	55.6	781	1	S64576	hypothetical prote
288	5	55.6	781	2	A86205	hypothetical prote
289	5	55.6	790	2	S27458	SM2 protein - yea
290	5	55.6	794	2	T37989	DNA mismatch repai
291	5	55.6	796	2	T03746	hypothetical prote
292	5	55.6	806	2	H70647	probable NADH dehy
293	5	55.6	808	1	S33708	nuclear steroid/th
294	5	55.6	811	2	PN0689	connectin I - chic
295	5	55.6	811	2	S76690	hypothetical prote
296	5	55.6	817	2	S51342	verprolin - yeast
297	5	55.6	823	1	A58788	procollagen C-endo
298	5	55.6	823	2	T02812	hypothetical prote
299	5	55.6	829	2	T33283	hypothetical prote
300	5	55.6	834	2	E69380	hypothetical prote
301	5	55.6	835	2	S11442	invasin - Yersinia
302	5	55.6	835	2	C86444	carbon-monoxide de
303	5	55.6	842	2	S49124	hypothetical prote
304	5	55.6	843	2	S33442	EF protein - strep
305	5	55.6	845	1	PRX51A	RNA-directed RNA p
306	5	55.6	866	1	JC4305	dynamitin II - huma
307	5	55.6	868	1	A36878	dynamitin 2 - rat
308	5	55.6	870	2	A53165	dynamitin II isoform
309	5	55.6	870	2	B53165	dynamitin II isoform
310	5	55.6	873	2	A47283	calphostin - fruit
311	5	55.6	899	2	S51341	SGP1 protein - yea
312	5	55.6	900	1	GNNMYV	genome polyprotein
313	5	55.6	920	2	T33263	cell division prot
314	5	55.6	929	2	G72677	hypothetical prote
315	5	55.6	929	2	G71420	hypothetical prote
316	5	55.6	930	2	A25923	progesterone recep
317	5	55.6	931	2	D86222	protein F7G19.9 li
318	5	55.6	933	1	ORHUP	progesterone recep
319	5	55.6	939	2	S28394	probable serine/th
320	5	55.6	940	2	T01834	hypothetical prote
321	5	55.6	951	2	T08987	probable cadmium-t
322	5	55.6	954	2	S57108	hypothetical prote
323	5	55.6	964	2	T15342	hypothetical prote
324	5	55.6	964	2	A75634	McRB-related prote
325	5	55.6	986	1	B58788	procollagen C-endo
326	5	55.6	990	2	I51618	nucleolar phospho
327	5	55.6	991	2	I49540	procollagen C-endo
328	5	55.6	1001	2	T00532	probable cadmium-t
329	5	55.6	1004	2	H70673	probable mbp prot
330	5	55.6	1009	1	DJBEM2	DNA-directed DNA p
331	5	55.6	1018	1	S73720	cytadherence acces
332	5	55.6	1027	1	GNLJST	pol polyprotein -
333	5	55.6	1047	2	T46489	hypothetical prote
334	5	55.6	1057	2	S45801	probable membrane
335	5	55.6	1070	2	T25836	hypothetical prote
336	5	55.6	1108	2	T38673	probable transcript
337	5	55.6	1113	2	H84105	hypothetical prote
338	5	55.6	1119	2	T50995	related to cytoske
339	5	55.6	1123	2	S36846	myosin-binding pro
340	5	55.6	1138	2	S24614	myosin-binding pro
341	5	55.6	1159	2	T43461	probable phosphodi
342	5	55.6	1171	1	QOKRFP	pyruvate (flavodox
343	5	55.6	1172	2	F70535	probable rpoB prot
344	5	55.6	1172	2	F84572	probable cadmium-t
345	5	55.6	1179	2	S31145	DNA-directed RNA p
346	5	55.6	1182	2	T29097	pro-pol-dutPase po
347	5	55.6	1199	2	S20969	Na+/Ca2+, K+-exchan
348	5	55.6	1211	2	T42230	AR4 protein - mous
349	5	55.6	1213	2	A54063	TAR4-binding prote
350	5	55.6	1217	2	T25894	hypothetical prote
351	5	55.6	1224	2	T42625	AF-4 protein - mou
352	5	55.6	1224	2	T40765	webl protein homol
353	5	55.6	1228	2	S56681	probable membrane
354	5	55.6	1240	2	T04193	hypothetical prote
355	5	55.6	1253	1	VHWV	structural polypyo
356	5	55.6	1254	2	JC7185	chromosome 1 Clorf
357	5	55.6	1268	2	B36502	insulin receptor-r
358	5	55.6	1268	2	T31420	C-terminal domain-
359	5	55.6	1276	2	T18526	SREBP cleavage act
360	5	55.6	1300	2	A36502	insulin receptor-r
361	5	55.6	1331	2	A72647	probable surface I
362	5	55.6	1337	2	A53824	nuclear pore membr
363	5	55.6	1341	2	T17285	hypothetical prote
364	5	55.6	1415	2	T21244	zys-9 protein - Ca
365	5	55.6	1418	2	T37264	phospholipase C (E
366	5	55.6	1473	2	A35186	salivary agglutini
367	5	55.6	1534	2	A56734	ritosome receptor,
368	5	55.6	1566	2	A43607	cell surface antiq
369	5	55.6	1582	2	T15308	hypothetical prote
370	5	55.6	1616	2	T47801	hypothetical prote
371	5	55.6	1627	1	IUYMAP	adhesin p1 precurs
372	5	55.6	1635	2	A41480	adhesin p1, group
373	5	55.6	1670	1	CGH03B	collagen alpha 3(I
374	5	55.6	1684	2	JW0057	gravin - human
375	5	55.6	1687	2	T30244	phosphodiesterase
376	5	55.6	1694	2	A83512	hypothetical prote
377	5	55.6	1706	2	T30175	exoribonuclease, v
378	5	55.6	1711	2	T31337	1,4-beta-glucanase
379	5	55.6	1719	2	T30174	exoribonuclease, v
380	5	55.6	1723	2	S58880	receptor DEC-205 -
381	5	55.6	1736	2	A47747	tight junction pro
382	5	55.6	1785	2	A45546	major merzoltie su
383	5	55.6	1794	2	T38459	hypothetical diver
384	5	55.6	1822	2	S33441	EF protein - strep
385	5	55.6	1829	2	T34239	hypothetical prote
386	5	55.6	1846	2	T10670	hypothetical prote
387	5	55.6	1963	2	T49914	callose synthase c
388	5	55.6	2144	2	S71490	asII protein - fru
389	5	55.6	2254	2	D86215	protein T6D22.14 l
390	5	55.6	2287	2	T21312	hypothetical prote
391	5	55.6	2290	1	GNNYE	genome polyprotein
392	5	55.6	2292	1	GNNYED	genome polyprotein
393	5	55.6	2292	1	GNNYEB	genome polyprotein
394	5	55.6	2292	2	S35961	capsid polyprotein

395	5	55.6	2292	2	S55401	capsid polypoteine	468	4	44.4	78	2	T00719	hypothetical prote
396	5	55.6	2447	2	T16870	hypothetical prote	469	4	44.4	79	2	T30118	hypothetical prote
397	5	55.6	2476	2	T34022	zonadhesin - pig	470	4	44.4	80	1	RHIDS	gonadoliberin I pr
398	5	55.6	2500	1	WMHUE2	HIV-EP2 enhancer-b	471	4	44.4	81	2	T07273	photosystem II pho
399	5	55.6	2796	2	JC4743	fatty-acid synthas	472	4	44.4	80	2	T29888	hypothetical prote
400	5	55.6	3131	2	S39842	emulatin synthetas	473	4	44.4	82	2	T29889	hypothetical prote
401	5	55.6	3133	2	S52093	hemocytin - silkw	474	4	44.4	82	2	T29893	hypothetical prote
402	5	55.6	3418	1	G02394	breast cancer tumo	475	4	44.4	82	2	T29891	hypothetical prote
403	5	55.6	4162	2	T42633	connectin/tltin -	476	4	44.4	82	2	T33088	hypothetical prote
404	5	55.6	4957	2	T03455	ALR protein - huma	477	4	44.4	84	2	A53012	serum response fac
405	5	55.6	5262	2	T03454	ALR protein - huma	478	4	44.4	84	2	T09540	polline rich prote
406	5	55.6	6839	2	S57242	twilchin [similari	479	4	44.4	87	1	FDFICP	antifreeze protein
407	5	55.6	7160	2	T27935	hypothetical prote	480	4	44.4	87	2	B31075	antifreeze protein
408	4	44.4	15	2	A40634	orf19 3' of eryk -	481	4	44.4	87	2	C31075	antifreeze protein
409	4	44.4	23	2	F61491	seed protein ws-6	482	4	44.4	87	2	T27141	hypothetical prote
410	4	44.4	25	2	T40692	cena protein (Igal	483	4	44.4	88	2	F85717	hypothetical prote
411	4	44.4	30	2	B60281	30k serine protein	484	4	44.4	88	2	A85743	hypothetical prote
412	4	44.4	32	2	B13897	alkaline phosphata	485	4	44.4	91	2	D86503	hypothetical prote
413	4	44.4	33	2	T22565	R-phycocerythrin ga	486	4	44.4	91	2	C72119	hypothetical prote
414	4	44.4	38	2	S40096	chlorophyll a/b-bl	487	4	44.4	92	2	T34744	hypothetical prote
415	4	44.4	41	2	B54357	muconate cyclolom	488	4	44.4	92	2	T51187	small zinc finger-
416	4	44.4	41	2	I68632	hypothetical myeli	489	4	44.4	93	2	T53642	insulin-like growt
417	4	44.4	43	2	H82712	hypothetical prote	490	4	44.4	93	2	A86491	CT001 hypothetical
418	4	44.4	43	2	T35464	hypothetical prote	491	4	44.4	93	2	T45593	small zinc finger-
419	4	44.4	47	2	E47395	histone H1 II-1 (c	492	4	44.4	93	2	F72130	conserved hypotet
420	4	44.4	48	1	T2A2R2	toxin RpiI - sea a	493	4	44.4	94	2	B84539	hypothetical prote
421	4	44.4	48	2	JN0355	neurotoxin I - sea	494	4	44.4	95	2	A56644	inverted repeat co
422	4	44.4	49	2	PC4007	hypothetical prote	495	4	44.4	96	2	T16179	hypothetical prote
423	4	44.4	50	2	B60718	phospholipase A2 h	496	4	44.4	98	2	T49562	alpha-1 type III c
424	4	44.4	51	2	S47401	outer membrane pro	497	4	44.4	99	2	C29826	hypothetical prote
425	4	44.4	52	2	A57336	ropA protein - Rhl	498	4	44.4	100	2	B72540	hypothetical prote
426	4	44.4	53	2	I53394	Soluble interleuk1	499	4	44.4	100	2	S65724	hypothetical prote
427	4	44.4	53	2	F47395	histone H1 II-1 (c	500	4	44.4	101	2	MNVUMB	extracellular hemo
428	4	44.4	53	2	B47395	histone H1 II-1 (c	501	4	44.4	101	1	MNVUBV	nonstructural prot
429	4	44.4	53	2	C47395	histone H1 II-1 (c	502	4	44.4	101	2	B72079	hypothetical prote
430	4	44.4	53	2	D47395	histone H1 II-1 (c	503	4	44.4	101	2	F86544	hypothetical prote
431	4	44.4	53	2	G47395	histone H1 II-2 (c	504	4	44.4	101	2	T19560	hypothetical prote
432	4	44.4	53	2	A72614	hypothetical prote	505	4	44.4	102	1	W7WLEP	E7 protein - deer
433	4	44.4	56	2	I56605	17 alpha-hydroxyla	506	4	44.4	102	1	PH1076	Ig light chain V r
434	4	44.4	60	2	S53364	ubiquinol--cytochr	507	4	44.4	102	2	C81682	conserved hypotet
435	4	44.4	62	1	CCBO17	hypothetical prote	508	4	44.4	102	2	F84709	probable glutaredo
436	4	44.4	62	2	T26847	antifreeze protein	509	4	44.4	102	2	F84709	Ig lambda chain C
437	4	44.4	63	2	B31075	antifreeze protein	510	4	44.4	103	2	B26167	conserved hypotet
438	4	44.4	64	2	D31075	hypothetical prote	511	4	44.4	103	2	G75513	gastirin precursor
439	4	44.4	65	2	S35024	yycd protein - Bac	512	4	44.4	104	1	GMBO	hypothetical 11.6k
440	4	44.4	66	2	S65971	hypothetical prote	513	4	44.4	104	2	J00863	Ig lambda chain C
441	4	44.4	66	2	PN0644	histone H1 I-1 (cl	514	4	44.4	105	1	L2HU	protein kinase (EC
442	4	44.4	67	2	I47395	hypothetical prote	515	4	44.4	105	2	S47430	hypothetical prote
443	4	44.4	67	2	G72372	archaeal histone -	516	4	44.4	105	2	A05188	spdd protein - str
444	4	44.4	67	2	A64321	archaeal histone -	517	4	44.4	105	2	T36213	nodc protein - Bra
445	4	44.4	67	2	A64457	archaeal histone -	518	4	44.4	105	2	S27493	R02D0.3 protein -
446	4	44.4	67	2	D64416	archaeal histone	519	4	44.4	105	2	S44838	outer membrane pro
447	4	44.4	67	2	D64513	lactocin S precurs	520	4	44.4	106	1	D82734	Ig kappa-2 chain C
448	4	44.4	68	1	A55457	synapotaegmin II p	521	4	44.4	106	1	K4RBBS	nonstructural prot
449	4	44.4	69	2	PC7068	cold shock DNA-bin	522	4	44.4	106	1	B48354	Ig heavy chain V r
450	4	44.4	69	2	A82489	6-phosphofructo-2-	523	4	44.4	106	2	S20774	Ig kappa-B4 chain
451	4	44.4	70	2	A4243	histone H1 I-1 (N-	524	4	44.4	106	2	G20907	hypothetical prote
452	4	44.4	70	2	A47395	ribosomal protein	525	4	44.4	106	2	G72632	hypothetical prote
453	4	44.4	71	2	T45384	hypothetical prote	526	4	44.4	106	2	S64649	RBL2 protein - yea
454	4	44.4	71	2	C82537	gene 35 protein -	527	4	44.4	107	2	T45621	hypothetical prote
455	4	44.4	72	2	S30980	hypothetical prote	528	4	44.4	107	2	B64845	hypothetical prote
456	4	44.4	72	2	T48971	conserved hypotet	529	4	44.4	107	2	A82653	protein-tyrosine-p-
457	4	44.4	73	2	H69035	hypothetical prote	530	4	44.4	108	2	D72662	hypothetical prote
458	4	44.4	73	2	T20911	hypothetical prote	531	4	44.4	108	2	F84479	En/Spw-like transp
459	4	44.4	73	2	E82788	orf13 protein - Mar	532	4	44.4	108	2	A72704	hypothetical prote
460	4	44.4	75	2	C44196	hypothetical prote	533	4	44.4	108	2	G72708	hypothetical prote
461	4	44.4	76	2	T17673	hypothetical prote	534	4	44.4	108	2	D71279	hypothetical prote
462	4	44.4	76	2	A75309	SNP2 protein - yea	535	4	44.4	109	2	T34639	hypothetical prote
463	4	44.4	76	2	I39960	hypothetical prote	536	4	44.4	109	2	T17325	probable PB protei
464	4	44.4	77	2	S48322	hypothetical prote	537	4	44.4	110	2	D70610	hypothetical prote
465	4	44.4	78	2	E82612	B1177_F2-71 protei	538	4	44.4	110	2	T46071	hypothetical prote
466	4	44.4	78	2	S72745	hypothetical prote	539	4	44.4	110	2	G72696	hypothetical prote
467	4	44.4	78	2	T35377	hypothetical prote	540	4	44.4	110	2	G72696	hypothetical prote

541	4	44.4	111	2	T16302	hypothetical prote	614	4	44.4	127	2	G83870	inner spore coat p
542	4	44.4	111	2	A81673	conserved hypotet	615	4	44.4	128	2	D72481	hypothetical prote
543	4	44.4	111	2	S17200	protein kinase (Ec	616	4	44.4	128	2	H84223	hypothetical prote
544	4	44.4	111	2	D72646	hypothetical prote	617	4	44.4	128	2	T05635	hypothetical prote
545	4	44.4	112	2	E82540	conserved hypotet	618	4	44.4	128	2	T32789	hypothetical prote
546	4	44.4	112	2	S30269	protein hdeb precu	619	4	44.4	129	2	B71350	probable ribosomal
547	4	44.4	112	2	F66023	hypothetical prote	620	4	44.4	130	1	R3MX8	probable ribosomal
548	4	44.4	112	2	A71515	hypothetical prote	621	4	44.4	130	2	S08079	Ig kappa chain pre
549	4	44.4	112	2	T01515	hypothetical prote	622	4	44.4	130	2	G70813	probable lpqs prot
550	4	44.4	112	2	C83938	hypothetical prote	623	4	44.4	130	2	D70966	probable phage pro
551	4	44.4	113	2	PC1281	N55 protein - hepa	624	4	44.4	130	2	T29887	hypothetical prote
552	4	44.4	113	2	S11640	hypothetical prote	625	4	44.4	130	2	D84363	hypothetical prote
553	4	44.4	113	2	D83220	hypothetical prote	626	4	44.4	130	2	T51879	hypothetical prote
554	4	44.4	113	2	T49533	hypothetical prote	627	4	44.4	130	2	A72553	hypothetical prote
555	4	44.4	113	2	B71088	hypothetical prote	628	4	44.4	131	2	T35341	hypothetical prote
556	4	44.4	114	2	S07898	endothelin 3 - tab	629	4	44.4	131	2	S60913	hypothetical prote
557	4	44.4	115	1	JN0318	guanylin precursor	630	4	44.4	131	2	H75320	hypothetical prote
558	4	44.4	115	1	KVMSL7	Ig kappa chain pre	631	4	44.4	131	2	E72649	hypothetical prote
559	4	44.4	115	2	E45213	Ca2+-transporting	632	4	44.4	131	2	T19134	hypothetical prote
560	4	44.4	115	2	G72568	hypothetical prote	633	4	44.4	131	2	E70920	probable moab prot
561	4	44.4	115	2	E70702	hypothetical prote	634	4	44.4	131	2	B67882	unknown protein F2
562	4	44.4	115	2	B84684	hypothetical prote	635	4	44.4	131	2	T24051	hypothetical prote
563	4	44.4	115	2	D84501	hypothetical prote	636	4	44.4	132	2	S49160	hypothetical prote
564	4	44.4	116	1	B46279	guanylin precursor	637	4	44.4	132	2	B70312	hypothetical prote
565	4	44.4	116	2	I51733	beta-2-microglobul	638	4	44.4	132	2	T17822	hypothetical prote
566	4	44.4	116	2	G70831	hypothetical prote	639	4	44.4	133	1	SXADP1	hexon-associated p
567	4	44.4	116	2	S70726	replication activa	640	4	44.4	133	2	S49487	long-chain-fatty-a
568	4	44.4	116	2	S70313	hypothetical prote	641	4	44.4	133	2	D69107	ribosomal protein
569	4	44.4	116	2	H72489	hypothetical prote	642	4	44.4	133	2	T31427	profilin - common
570	4	44.4	117	2	T02872	probable lipid tra	643	4	44.4	133	2	G75432	hypothetical prote
571	4	44.4	117	2	S21668	Ig kappa chain v r	644	4	44.4	134	2	B37476	E3 gene encoding h
572	4	44.4	117	2	B82614	hypothetical prote	645	4	44.4	134	2	JC4772	hypothetical prote
573	4	44.4	117	2	E75537	conserved hypotet	646	4	44.4	134	2	S74701	hypothetical prote
574	4	44.4	118	2	S12440	Ig lambda chain (M	647	4	44.4	134	2	A83296	hypothetical prote
575	4	44.4	118	2	S12441	Ig lambda chain (K	648	4	44.4	134	2	E75416	hypothetical prote
576	4	44.4	118	2	S12442	Ig lambda chain (K	649	4	44.4	135	2	A48491	hypothetical prote
577	4	44.4	118	2	D72000	hypothetical prote	650	4	44.4	135	2	C64453	translating moltilty
578	4	44.4	118	2	T35747	hypothetical prote	651	4	44.4	135	2	T31260	translacion inflita
579	4	44.4	118	2	T07905	low-carbon dioxide	652	4	44.4	135	2	G72714	hypothetical prote
580	4	44.4	118	2	G66623	hypothetical prote	653	4	44.4	136	2	T35335	hypothetical prote
581	4	44.4	118	2	S34346	hypothetical prote	654	4	44.4	137	2	T06390	histone H2B-3 - to
582	4	44.4	119	2	S08077	Ig kappa chain pre	655	4	44.4	137	2	B28534	lucC 5'-region hyp
583	4	44.4	119	2	E72526	hypothetical prote	656	4	44.4	138	2	B45244	transcription regu
584	4	44.4	120	2	S21667	Ig kappa chain v r	657	4	44.4	138	2	T04560	hypothetical prote
585	4	44.4	120	2	S21666	Ig kappa chain v r	658	4	44.4	138	2	B83366	hypothetical prote
586	4	44.4	120	2	T23978	hypothetical prote	659	4	44.4	138	2	G85806	hypothetical prote
587	4	44.4	120	2	G86402	hypothetical prote	660	4	44.4	138	2	T29558	hypothetical prote
588	4	44.4	120	2	A72539	histidine triad-li	661	4	44.4	139	2	T34633	hypothetical prote
589	4	44.4	121	2	E70313	ribosomal phosphop	662	4	44.4	139	2	T23471	hypothetical prote
590	4	44.4	121	2	JC4622	hypothetical prote	663	4	44.4	140	2	H64664	conserved hypotet
591	4	44.4	121	2	T32888	hypothetical prote	664	4	44.4	140	2	JU0041	hypothetical 14.5K
592	4	44.4	121	2	S75660	hypothetical prote	665	4	44.4	141	1	S34067	H+-transporting AT
593	4	44.4	122	2	T22096	hypothetical prote	666	4	44.4	141	2	T36028	probable iron sulf
594	4	44.4	123	2	G71143	hypothetical prote	667	4	44.4	142	1	S42722	interleukin-3 prec
595	4	44.4	123	2	H75059	hypothetical prote	668	4	44.4	142	1	S42721	interleukin-3 prec
596	4	44.4	123	2	B75546	hypothetical prote	669	4	44.4	142	2	D69510	oxaloacetate decar
597	4	44.4	123	2	C71423	hypothetical prote	670	4	44.4	142	2	A53624	microtubule associ
598	4	44.4	123	2	T20279	hypothetical prote	671	4	44.4	142	2	T10483	rega-associated pr
599	4	44.4	124	1	MMIHMS	nonstructural prote	672	4	44.4	142	2	A70757	hypothetical prote
600	4	44.4	124	2	H72263	conserved hypotet	673	4	44.4	142	2	E83975	hypothetical prote
601	4	44.4	124	2	T36292	hypothetical prote	674	4	44.4	143	1	S04230	H+-transporting AT
602	4	44.4	125	2	S78697	probable export pr	675	4	44.4	143	1	H65243	yJHL protein - Esc
603	4	44.4	125	2	C70977	hypothetical prote	676	4	44.4	143	2	S34065	H+-transporting AT
604	4	44.4	125	2	B72610	hypothetical prote	677	4	44.4	143	2	T15722	hypothetical prote
605	4	44.4	125	2	T38537	probable single-st	678	4	44.4	144	1	SXAD12	hexon-associated p
606	4	44.4	126	2	T24181	hypothetical prote	679	4	44.4	144	2	S66570	biotin carboxyl ca
607	4	44.4	126	2	H75433	hypothetical prote	680	4	44.4	144	2	PC1110	riboflavin synhas
608	4	44.4	126	2	G84235	hypothetical prote	681	4	44.4	144	2	C75376	hypothetical prote
609	4	44.4	127	1	C64140	probable glycyI ra	682	4	44.4	144	2	JC2102	lectin-related pro
610	4	44.4	127	2	E72263	flagellar basal-bo	683	4	44.4	144	2	T46334	hypothetical prote
611	4	44.4	127	2	S00653	chlorophyll a/b-bi	684	4	44.4	144	2	B83069	hypothetical prote
612	4	44.4	127	2	C85091	arabinogalactan-pr	685	4	44.4	144	2	A57362	gyrb protein - Str
613	4	44.4	127	2	F64039	hypothetical prote	686	4	44.4	145	2	A82619	pilus protein Xf19

687	4	44.4	145	2	C70328	conserved hypot	760	4	44.4	160	2	F84350	hypothetical prote
688	4	44.4	145	2	S63134	probable membrane	761	4	44.4	161	1	AFKTA	allophycocyanin al
689	4	44.4	145	2	G70424	pillin - Aquifex ae	762	4	44.4	161	2	S44757	ribosomal protein
690	4	44.4	146	1	HBTR1	hemoglobin I beta	763	4	44.4	161	2	S43260	helix-loop-helix p
691	4	44.4	146	2	T03268	histone H2B1 - com	764	4	44.4	161	2	C36810	hypothetical prote
692	4	44.4	146	2	S76168	hypothetical prote	765	4	44.4	161	2	H82782	hypothetical prote
693	4	44.4	146	2	G73544	hypothetical prote	766	4	44.4	161	2	T37105	probable integrat
694	4	44.4	146	2	F82418	cytochrome c554 VC	767	4	44.4	162	2	H69504	hypothetical prote
695	4	44.4	147	2	T33327	hypothetical prote	768	4	44.4	162	2	C72699	hypothetical prote
696	4	44.4	147	2	T49624	hypothetical prote	769	4	44.4	162	2	G96671	hypothetical prote
697	4	44.4	147	2	T02273	hypothetical prote	770	4	44.4	162	2	C72655	hypothetical prote
698	4	44.4	148	2	G83613	probable acetyltra	771	4	44.4	163	1	YC6C3H	hypothetical prote
699	4	44.4	148	2	C84690	probable membrane	772	4	44.4	163	2	S15940	acetylactate synth
700	4	44.4	148	2	T21488	hypothetical prote	773	4	44.4	163	2	B64131	acetylactate synth
701	4	44.4	148	2	B96721	probable thioredox	774	4	44.4	163	2	B85490	hypothetical prote
702	4	44.4	149	2	C72611	hypothetical prote	775	4	44.4	163	2	F83059	acetylactate synth
703	4	44.4	149	2	F75509	hypothetical prote	776	4	44.4	163	2	H81066	acetylactate synth
704	4	44.4	150	2	G72660	hypothetical prote	777	4	44.4	164	1	TVBE11	transferring prote
705	4	44.4	150	2	G84173	archaeal histone A	778	4	44.4	164	2	H71850	flavodoxin - Helic
706	4	44.4	151	1	S76566	heat shock protein	779	4	44.4	164	2	A64665	flavodoxin - Helic
707	4	44.4	151	2	T09256	hypothetical prote	780	4	44.4	164	2	T38684	mitochondrial inne
708	4	44.4	151	2	F82427	hypothetical prote	781	4	44.4	164	2	S35021	nolB protein - Rhl
709	4	44.4	151	2	B45696	23-28K immunoreact	782	4	44.4	164	2	B70922	hypothetical prote
710	4	44.4	151	2	D72706	hypothetical prote	783	4	44.4	164	2	S44659	ZK351.3 protein -
711	4	44.4	151	2	S33205	proteinase 5 - but	784	4	44.4	164	2	S18038	homeotic protein S
712	4	44.4	152	2	S08350	superoxide dismuta	785	4	44.4	164	2	JC4675	transcripton fact
713	4	44.4	152	2	B71119	probable aspartate	786	4	44.4	165	2	A31635	neural cell adhesi
714	4	44.4	152	2	S57422	cysteine proteinas	787	4	44.4	165	2	T39187	probable arylalkyl
715	4	44.4	152	2	T09253	heat shock protein	788	4	44.4	165	2	T47011	hypothetical prote
716	4	44.4	152	2	T30753	hypothetical prote	789	4	44.4	166	2	H83630	hypothetical prote
717	4	44.4	152	2	T31054	hypothetical prote	790	4	44.4	167	1	S32716	ribosome releasing
718	4	44.4	152	2	A84303	hypothetical prote	791	4	44.4	167	1	UYPWF3	nonstructural prot
719	4	44.4	152	2	T08401	hypothetical prote	792	4	44.4	167	2	S73043	hypothetical prote
720	4	44.4	152	2	T494126	hypothetical prote	793	4	44.4	168	2	S67993	amylogenin - maize
721	4	44.4	152	2	T11658	mitochondrial impo	794	4	44.4	168	2	S52994	arabinogalactan-11
722	4	44.4	152	2	S35206	serine proteinase	795	4	44.4	168	2	S60970	hypothetical prote
723	4	44.4	153	2	B49854	probable reca regu	796	4	44.4	169	2	T30684	probable dual spec
724	4	44.4	153	2	F83192	conserved hypot	797	4	44.4	169	2	T29283	peptidylprolyl iso
725	4	44.4	153	2	S29377	hypothetical prote	798	4	44.4	169	2	H72611	hypothetical prote
726	4	44.4	153	2	E72506	hypothetical prote	799	4	44.4	169	2	D72347	hypothetical prote
727	4	44.4	153	2	S73143	photosystem I chal	800	4	44.4	170	2	T35957	hypothetical prote
728	4	44.4	153	2	JC5854	polyketide synthas	801	4	44.4	171	2	A45562	mezoquite surface
729	4	44.4	154	2	T28434	transcription regu	802	4	44.4	171	2	H69397	conserved hypot
730	4	44.4	154	2	B69105	hypothetical prote	803	4	44.4	171	2	T41803	AcMNPV orf58 - Bom
731	4	44.4	154	2	A85551	hypothetical prote	804	4	44.4	171	2	S41642	hypothetical prote
732	4	44.4	154	2	PC4031	hypothetical 154 p	805	4	44.4	171	2	H70527	probable PE protei
733	4	44.4	154	2	S55017	hypothetical prote	806	4	44.4	171	2	T37090	probable secreted
734	4	44.4	154	2	F72741	hypothetical prote	807	4	44.4	172	2	T21753	hypothetical prote
735	4	44.4	155	1	SNFR2	pathogenesis-relat	808	4	44.4	173	1	RUPSEO	tubredoxin II - Ps
736	4	44.4	155	2	T11670	pathogenesis relat	809	4	44.4	173	2	T39632	peptidylprolyl iso
737	4	44.4	155	2	E64105	acetyl-CoA carboxy	810	4	44.4	174	2	F70674	hypothetical prote
738	4	44.4	155	2	H86721	hypothetical prote	811	4	44.4	174	2	A45356	neuromedin U precu
739	4	44.4	155	2	A81065	sigma-54 dependant	812	4	44.4	175	2	C83141	inorganic pyrophos
740	4	44.4	155	2	H75302	conserved hypot	813	4	44.4	175	2	S49992	ribulose-1,5-bisph
741	4	44.4	156	1	SNFR1	pathogenesis-relat	814	4	44.4	175	2	E56271	long polar flindria
742	4	44.4	156	2	C84027	molybdoplerin conv	815	4	44.4	175	2	S71560	early light-induce
743	4	44.4	156	2	D81433	probable integrat	816	4	44.4	175	2	T39781	hypothetical prote
744	4	44.4	156	2	D70541	hypothetical prote	817	4	44.4	175	2	D75514	hypothetical prote
745	4	44.4	157	2	S04939	heat shock 22K pro	818	4	44.4	176	2	A26882	pilZ hypot
746	4	44.4	157	2	H72751	hypothetical prote	819	4	44.4	176	2	D45213	Ca2+-transporting
747	4	44.4	157	2	S31634	hypothetical prote	820	4	44.4	176	2	I53630	cell surface glyco
748	4	44.4	157	2	F83464	flagellar motor sw	821	4	44.4	176	2	D71434	hypothetical prote
749	4	44.4	157	2	S49793	hypothetical prote	822	4	44.4	176	2	T45308	hypothetical prote
750	4	44.4	157	2	T25638	hypothetical prote	823	4	44.4	176	2	S02210	con-8 protein - Ne
751	4	44.4	158	2	S20518	hypothetical prote	824	4	44.4	176	2	F72651	hypothetical prote
752	4	44.4	158	2	S20517	hypothetical prote	825	4	44.4	177	2	T06306	hypothetical prote
753	4	44.4	158	2	T30379	biotin carboxyl ca	826	4	44.4	177	2	D83111	transcription anti
754	4	44.4	158	2	D64365	hypothetical prote	827	4	44.4	177	2	D81392	probable transcrip
755	4	44.4	159	2	H64778	ybak protein - Esc	828	4	44.4	177	2	A83064	hypothetical prote
756	4	44.4	159	2	B85546	hypothetical prote	829	4	44.4	177	2	D82638	hypothetical prote
757	4	44.4	159	2	T48837	hypothetical prote	830	4	44.4	178	1	E69389	hypothetical prote
758	4	44.4	160	2	S78308	plascocquinoi--plas	831	4	44.4	178	1	T09585	high mobility grou
759	4	44.4	160	2	H69015	ribosomal protein	832	4	44.4	178	2	T09584	high mobility grou



833	4	44.4	178	2	S24298	906	4	44.4	191	2	T18828	hypothetical prote
834	4	44.4	178	2	T45019	907	4	44.4	191	2	S28736	hypothetical prote
835	4	44.4	178	2	B75406	908	4	44.4	191	2	J01099	hypothetical 21.5k
836	4	44.4	179	1	A42840	909	4	44.4	191	2	H82047	conserved hypotnet
837	4	44.4	179	2	T04897	910	4	44.4	192	2	S15930	hypothetical 21.5k
838	4	44.4	179	2	T22734	911	4	44.4	192	2	S55252	mushroom-inducing
839	4	44.4	180	1	IGHU2	912	4	44.4	193	2	A69155	hypothetical prote
840	4	44.4	180	1	HLHUC3	913	4	44.4	193	2	S74058	hypothetical prote
841	4	44.4	180	2	G70836	914	4	44.4	193	2	E96766	hypothetical prote
842	4	44.4	180	2	G29504	915	4	44.4	193	2	S38467	hypothetical prote
843	4	44.4	180	2	T29803	916	4	44.4	194	2	T16382	hypothetical prote
844	4	44.4	181	2	B60738	917	4	44.4	194	2	H72037	conserved hypotnet
845	4	44.4	181	2	D49233	918	4	44.4	194	2	C86586	Cr647 hypothetical
846	4	44.4	181	2	F84934	919	4	44.4	194	2	E75297	hypothetical prote
847	4	44.4	181	2	C81041	920	4	44.4	195	1	S74599	2-amino-4-hydroxy-
848	4	44.4	181	2	D72060	921	4	44.4	195	1	Q0ECAA	yaca protein - Esc
849	4	44.4	181	2	A86564	922	4	44.4	195	1	B32048	gliding motility p
850	4	44.4	181	2	T49728	923	4	44.4	195	2	E85492	hypothetical prote
851	4	44.4	181	2	A70673	924	4	44.4	195	2	A83902	hypothetical prote
852	4	44.4	181	2	T04643	925	4	44.4	196	2	PC1148	hypothetical prote
853	4	44.4	181	2	B84162	926	4	44.4	196	2	JC2391	amelogenin precurs
854	4	44.4	181	2	E86488	927	4	44.4	196	2	C70591	hypothetical prote
855	4	44.4	181	2	A86451	928	4	44.4	196	2	G85435	hypothetical prote
856	4	44.4	181	2	F82750	929	4	44.4	196	2	T13595	hypothetical prote
857	4	44.4	182	2	H70818	930	4	44.4	197	2	S57948	HMG1/Y protein - g
858	4	44.4	182	2	A82684	931	4	44.4	197	2	A49247	merozoite surface
859	4	44.4	182	2	T35142	932	4	44.4	197	2	C49247	merozoite surface
860	4	44.4	182	2	S73046	933	4	44.4	197	2	PQ0327	heparin-binding fi
861	4	44.4	182	2	JC7196	934	4	44.4	198	2	D82253	dedd protein VC100
862	4	44.4	182	2	T16472	935	4	44.4	198	2	E70357	conserved hypotnet
863	4	44.4	183	2	S02423	936	4	44.4	198	2	A96811	unknown protein T1
864	4	44.4	183	2	I67610	937	4	44.4	199	2	P83540	probable alkyl hyd
865	4	44.4	183	2	A32338	938	4	44.4	199	2	H75401	hypothetical prote
866	4	44.4	183	2	I54288	939	4	44.4	199	2	T24595	hypothetical prote
867	4	44.4	183	2	I79496	940	4	44.4	200	2	S55609	hypothetical prote
868	4	44.4	183	2	S72802	941	4	44.4	200	2	T37187	hypothetical prote
869	4	44.4	183	2	T34330	942	4	44.4	200	2	T50442	hypothetical prote
870	4	44.4	183	2	G83379	943	4	44.4	201	2	T06229	probable superoxid
871	4	44.4	183	2	S24960	944	4	44.4	201	2	T06800	superoxide dismuta
872	4	44.4	184	2	A70149	945	4	44.4	201	2	T46670	probable 3,5-epime
873	4	44.4	184	2	E82345	946	4	44.4	201	2	D38625	GRP-binding protei
874	4	44.4	184	2	C96506	947	4	44.4	201	2	S06147	GRP-binding protei
875	4	44.4	184	2	T00431	948	4	44.4	201	2	JC5337	GRP-binding protei
876	4	44.4	184	2	S77928	949	4	44.4	201	2	T09047	hypothetical protei
877	4	44.4	185	2	D75602	950	4	44.4	202	1	HTTVH2	hemorrhagic protei
878	4	44.4	185	2	S76706	951	4	44.4	202	1	JQ1556	coat protein - Ery
879	4	44.4	185	2	A82636	952	4	44.4	202	2	T15874	hypothetical prote
880	4	44.4	186	2	S43174	953	4	44.4	202	2	S26074	hypothetical prote
881	4	44.4	186	2	F83753	954	4	44.4	202	2	A42410	rydocan precursor
882	4	44.4	186	2	S74835	955	4	44.4	202	2	T47319	hypothetical prote
883	4	44.4	186	2	T31951	956	4	44.4	202	2	F84382	hypothetical prote
884	4	44.4	186	2	E82625	957	4	44.4	202	2	A61379	22k antigen - Rick
885	4	44.4	187	1	B39830	958	4	44.4	203	2	B75300	hypothetical prote
886	4	44.4	187	2	T36331	959	4	44.4	203	2	JC4990	chymotrypsin inh
887	4	44.4	187	2	T34834	960	4	44.4	204	2	G82149	Holliday junction
888	4	44.4	187	2	F65009	961	4	44.4	204	2	T48091	hypothetical prote
889	4	44.4	187	2	S76728	962	4	44.4	204	2	G81876	probable membrane
890	4	44.4	187	2	H84100	963	4	44.4	204	2	D84190	tryptophan synthas
891	4	44.4	188	2	G70892	964	4	44.4	205	1	TVHUYF	GRP-binding protei
892	4	44.4	188	2	A75392	965	4	44.4	205	1	TVDGYP	GRP-binding protei
893	4	44.4	189	2	S52907	966	4	44.4	205	1	TVMSPV	GRP-binding protei
894	4	44.4	189	2	C83134	967	4	44.4	205	1	TYRTVP	GRP-binding protei
895	4	44.4	189	2	C82515	968	4	44.4	205	2	S38339	GRP-binding protei
896	4	44.4	190	2	S25740	969	4	44.4	205	2	T33117	hypothetical prote
897	4	44.4	190	2	E64775	970	4	44.4	205	2	T27923	hypothetical prote
898	4	44.4	190	2	G85542	971	4	44.4	205	2	T34724	probable membrane
899	4	44.4	190	2	T32265	972	4	44.4	206	1	B69128	precursorin isomer
900	4	44.4	190	2	T00735	973	4	44.4	206	2	I68547	PA338 protein - hu
901	4	44.4	190	2	T40866	974	4	44.4	206	2	JL0058	H-2 class I histoc
902	4	44.4	191	2	G71835	975	4	44.4	206	2	JL0059	H-2 class I histoc
903	4	44.4	191	2	G64684	976	4	44.4	206	2	S44894	Zk1236.6 protein -
904	4	44.4	191	2	S76377	977	4	44.4	206	2	T49138	hypothetical prote
905	4	44.4	191	2	C41892	978	4	44.4	206	2	D72323	conserved hypotnet

```

979 4 44.4 206 2 T27766 hypothetical prote
980 4 44.4 206 2 B82563 hypothetical prote
981 4 44.4 207 2 H72550 probable inorganic
982 4 44.4 207 2 S77728 type 4 fibrinai bl
983 4 44.4 207 2 E69459 hypothetical prote
984 4 44.4 207 2 B75327 hypothetical prote
985 4 44.4 207 2 A83540 phospholipase acce
986 4 44.4 207 2 T08109 oleosin-like prote
987 4 44.4 208 2 T09216 interluekin-6 prec
988 4 44.4 208 2 A26496 granulocyte colony
989 4 44.4 208 2 B49444 ig lambda chain (N
990 4 44.4 208 2 C34503 small nuclear ribo
991 4 44.4 208 2 T46927 hypothetical prote
992 4 44.4 208 2 T34763 probable transcrip
993 4 44.4 208 2 E72241 hypothetical prote
994 4 44.4 208 2 S73699 adhesin PI precurs
995 4 44.4 208 2 T08132 oleosin-like prote
996 4 44.4 208 2 A43596 Spec3 protein - se
997 4 44.4 209 2 S44650 f42hi0.3 protein -
998 4 44.4 210 1 S03583 ribosomal protein
999 4 44.4 210 1 B64416 probable precorrin
1000 4 44.4 210 2 D84415 30S ribosomal prot

```

## ALIGNMENTS

```

RESULT 1
A:1487
protein p60 precursor - listeria monocytogenes
N:Alternate names: Invasion-associated protein
C:Species: Listeria monocytogenes
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1996
C:Accession: A41487; B41487
R:Koehler, S.; Leimister-Maehtler, M.; Chakraborty, T.; Lottspiech, F.; Goebel, W.
Infect. Immun. 59, 1943-1950, 1990
A:Title: The gene coding for protein p60 of Listeria monocytogenes and its use as a spec
A:Reference number: A41487; MUID:90256283
A:Accession: A41487
A:Molecule type: DNA
A:Residues: 1-484 <KOE>
A:Cross-references: GB:X52268
A:Accession: B41487
A:Molecule type: protein
A:Residues: 28-49 <KO2>
C:Genetics:
A:Gene: iap
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-464/Product: protein p60 #status predicted <MAT>

```

```

Query Match          100.0%; Score 9; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VSTPVAPPTQ 9
Db 145 VSTPVAPPTQ 153

```

```

RESULT 2
F83231
hypothetical protein PA3307 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83231
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337
A:Accession: F83231

```

```

A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-102 <STO>
A:Cross-references: GB:AE004753; GB:AE004091; NID:99949433; PIDN:AMG06695.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3307

```

```

Query Match          66.7%; Score 6; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 PVAPPTQ 9
Db 27 PVAPPTQ 32

```

```

RESULT 3
S35162
STR-21 protein - potato
C:Species: Solanum tuberosum (potato)
C>Date: 13-Jan-1995 #sequence_revision 09-May-1997 #text_change 20-Aug-1999
C:Accession: S35162; S11869
R:Matton, D.P.; Prescott, G.; Bertrand, C.; Camirand, A.; Brisson, N.
Plant Mol. Biol. 22, 279-291, 1993
A:Title: Identification of cis-acting elements involved in the regulation of the path
A:Reference number: S35161; MUID:93283632
A:Accession: S35162
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-155 <MAT>
A:Cross-references: EMBL:M29042; NID:9169577; PIDN:AAA02829.1; PID:9169578
C:Genetics:
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1990
A:Gene: STR-21
A:Introns: 57/1
C:Superfamily: pathogenesis-related protein

```

```

Query Match          66.7%; Score 6; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3 TPVAPT 8
Db 12 TPVAPT 17

```

```

RESULT 4
JC4806
core protein G - phage phi-K
C:Species: phage phi-K
C>Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 20-Jun-2000
C:Accession: JC4806; B04253; A04253
R:Kodaira, K.; Oki, M.; Kakikawa, M.; Kimoto, H.; Taketo, A.
J. Biochem. 119, 1062-1069, 1996
A:Title: The virion proteins encoded by bacteriophage phi-K and its host-range mutant
A:Reference number: JC4804; MUID:96424967
A:Accession: JC4806
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-187 <KOD>
A:Cross-references: EMBL:X60323; NID:91478118; PIDN:CAA42892.1; PID:91478127
R:Stms, J.; Capon, D.; Dressler, D.
J. Biol. Chem. 254, 12615-12628, 1979
A:Title: dnag (primase)-dependent origins of DNA replication. Nucleotide sequences of
A:Reference number: A92247; MUID:80049950
A:Accession: B04253
A:Molecule type: DNA
A:Residues: 165-187 <SIM>
A:Comment: This protein is one of the structural components of the bacteriophage caps
C:Genetics:

```

A:Gene: G  
C:Superfamily: phage phi-X174 gene G protein  
C:Keywords: capsid protein

Query Match 66.7%; Score 6; DB 2; Length 187;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPT 8  
|||||  
DB 25 TPVAPT 30

RESULT 5  
C86847  
transcription regulator Tena [imported] - Lactococcus lactis subsp. lactis (strain IL140  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C86847  
R:Boletín, A.; Wincker, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich  
Genome Res. In press, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium.  
A:Reference number: A86625  
A:Accession: C86847  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-218 <STO>  
A:Cross-references: GB:AE005176; NID:q12724803; PIDN:AAK05877.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: tena  
C:Superfamily: transcription activator tena

Query Match 66.7%; Score 6; DB 2; Length 218;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPT 8  
|||||  
DB 103 TPVAPT 108

RESULT 6  
F71504  
hypothetical protein CT504 - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 18-Aug-2000  
C:Accession: F71504  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trac  
A:Reference number: A71570; MUID:99000809  
A:Accession: F71504  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-288 <ARN>  
A:Cross-references: GB:AE001323; GB:AE001273; NID:g3328931; PIDN:AAC68105.1; PID:g332894  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: CT504  
C:Superfamily: conserved hypothetical protein TC0791

Query Match 66.7%; Score 6; DB 2; Length 288;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVAP 7  
|||||  
DB 131 STPVAP 136

RESULT 7

F75466  
hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: F75466  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: F75466  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-293 <WHI>  
A:Cross-references: GB:AE001940; GB:AE000513; NID:g6458577; PIDN:AAF10447.1; PID:g645  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0864  
A:Map position: 1

Query Match 66.7%; Score 6; DB 2; Length 293;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVAP 7  
|||||  
DB 227 STPVAP 232

RESULT 8  
T34992  
probable lipoprotein - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T34992  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1998  
A:Reference number: Z21550  
A:Accession: T34992  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-447 <OLI>  
A:Cross-references: EMBL:AL031182; PIDN:CAA20169.1; GSPDB:GN00070; SCOEDB:SCA42.17C  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SCA42.17C

Query Match 66.7%; Score 6; DB 2; Length 447;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPT 8  
|||||  
DB 63 TPVAPT 68

RESULT 9  
F64623  
amidase - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: F64623  
R:Tombo, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khatak, H.G.; Glodek, A.; McKe  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney,  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467  
A:Accession: F64623  
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-453 <TOM>  
A:Cross-references: GB:AE000594; GB:AE000511; NID:g2313957; PIDN:AAD07880.1; PID:g231396  
C:Superfamily: Indoleacetamide hydrolase

Query Match 66.7%; Score 6; DB 1; Length 453;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPT 8  
|||||  
Db 373 TPVAPT 378

RESULT 10  
A71891  
glu-tRNA amidotransferase, chain A - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C>Date: 12-Feb-1999 #sequence-revision 12-Feb-1999 #text-change 22-Jun-1999  
C:Accession: A71891  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A:Reference number: A71800; MUID:99120557  
A:Accession: A71891  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-453 <ARN>  
A:Cross-references: GB:AE001507; GB:AE001439; NID:g4155326; PIDN:AAD06348.1; PID:g415533  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: gata  
C:Superfamily: Indoleacetamide hydrolase

Query Match 66.7%; Score 6; DB 2; Length 453;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPT 8  
|||||  
Db 373 TPVAPT 378

RESULT 11  
T23087  
hypothetical protein H13N06.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence-revision 15-Oct-1999 #text-change 18-Feb-2000  
C:Accession: T23087  
R:Lennard, N.  
submitted to the EMBL Data Library, October 1997  
A:Reference number: Z19673  
A:Accession: T23087  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-460 <WIL>  
A:Cross-references: EMBL:Z99942; PIDN:CAB17068.1; GSPDB:GN00028; CESP:H13N06.2  
C:Genetics:  
A:Gene: CESP:H13N06.2  
A:Map position: X  
A:Introns: 17/1; 294/1; 331/3; 364/1; 414/2; 432/2

Query Match 66.7%; Score 6; DB 2; Length 460;  
Best Local Similarity 100.0%; Pred. No. 28;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 3 TPVAPT 8  
|||||  
Db 73 TPVAPT 78

RESULT 12  
S66953  
hypothetical protein YOR070c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein O2931  
C:Species: Saccharomyces cerevisiae  
C>Date: 12-Jul-1996 #sequence-revision 12-Jul-1996 #text-change 05-Nov-1999  
C:Accession: S66953  
R:Bohn, C.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Dang, D.V.; Valens, M.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66929  
A:Accession: S66953  
A:Molecule type: DNA  
A:Residues: 1-637 <BOH>  
A:Cross-references: EMBL:Z74978; NID:g1420220; PID:e251984; PID:g1420221; GSPDB:GN000  
C:Genetics:  
A:Gene: MIPS:YOR070c  
A:Map position: 15R

Query Match 66.7%; Score 6; DB 2; Length 637;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPT 8  
|||||  
Db 181 TPVAPT 186

RESULT 13  
T49633  
glucan 1,4-alpha-glucosidase related protein [imported] - Neurospora crassa  
N:Alternate names: protein B5022.150  
C:Species: Neurospora crassa  
C>Date: 02-Jun-2000 #sequence-revision 02-Jun-2000 #text-change 02-Jun-2000  
C:Accession: T49633  
R:Schulte, U.; Aign, V.; Hohnsels, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu  
submitted to the Protein Sequence Database, May 2000  
A:Reference number: Z25022  
A:Accession: T49633  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-748 <SCH>  
A:Cross-references: EMBL:AL355932; GSPDB:GN00116; NCSP:B5022.150  
C:Genetics:  
A:Gene: NCSP:B5022.150  
A:Map position: 6  
A:Introns: 701/1

Query Match 66.7%; Score 6; DB 2; Length 748;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTRPVA 6  
|||||  
Db 506 VSTRPVA 511

RESULT 14  
H84913  
probable ATP-dependent RNA helicase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Feb-2001 #sequence-revision 02-Feb-2001 #text-change 02-Feb-2001  
C:Accession: H84913

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.: Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, A.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487  
A:Accession: H84913  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-748 <STO>  
A:Cross-references: GB:AE002093; NID:g2275211; PIDN:AAB63833.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g47330  
A:Map position: 2

Query Match 66.7%; Score 6; DB 2; Length 748;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPTO 9  
|||||  
DB 725 PVAPTO 730

RESULT 15  
A47282  
calcium-binding protein calphotin - fruit fly (*Drosophila melanogaster*)  
C:Species: *Drosophila melanogaster*  
C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
C:Accession: A47282  
R:Martin, J.H.; Benzer, S.; Rudnicka, M.; Miller, C.A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 1531-1535, 1993  
A:Title: Calphotin: a *Drosophila* photoreceptor cell calcium-binding protein.  
A:Reference number: A47282; MUID:93165729  
A:Accession: A47282  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-865 <MAR>  
A:Cross-references: GB:L02111; NID:g157031; PIDN:AAA28405.1; PID:g157032  
A:Experimental source: photoreceptor cells  
A:Note: sequence extracted from NCBI backbone (NCBIN:124955, NCBI:124956)  
C:Genetics:  
A:Gene: FlyBase:Cpn  
A:Cross-references: FlyBase:FBgn0010218  
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
C:Keywords: calcium binding

Query Match 66.7%; Score 6; DB 2; Length 865;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVA 6  
|||||  
DB 97 VSTPVA 102

RESULT 16  
T50652  
AP-3 complex beta3a chain [imported] - human  
C:Species: *Homo sapiens* (man)  
C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
C:Accession: T50652  
R:Simpson, F.; Peden, A.A.; Christopoulos, L.; Robinson, M.S.  
J. Cell Biol. 137, 835-845, 1997  
A:Title: Characterization of the adaptor-related protein complex, AP-3.  
A:Reference number: 225159; MUID:97296491  
A:Accession: T50652  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1093 <SIM>  
A:Cross-references: EMBL:U91931; PIDN:AAD03778.1

Query Match 66.7%; Score 6; DB 2; Length 1093;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVA 6  
|||||  
DB 826 VSTPVA 831

RESULT 17  
T50651  
AP3-complex beta-3a-adaptin chain [imported] - human  
C:Species: *Homo sapiens* (man)  
C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
C:Accession: T50651  
R:Deil/Angelica, E.C.; Ooi, C.E.; Bonifacio, J.S.  
J. Biol. Chem. 272, 15078-15084, 1997  
A:Title: Beta3a-adaptin, a subunit of the adaptor-like complex AP-3.  
A:Reference number: 225158; MUID:97326075  
A:Accession: T50651  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1094 <DEL>  
A:Cross-references: EMBL:U81504; PIDN:AAB61638.1  
C:Function:  
A:Description: involved in the regulation of intracellular protein trafficking  
C:Keywords: heterotrimer; phosphoprotein

Query Match 66.7%; Score 6; DB 2; Length 1094;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVA 6  
|||||  
DB 827 VSTPVA 832

RESULT 18  
T18295  
AP-3 adaptor complex beta3a chain - mouse  
C:Species: *Mus musculus* (house mouse)  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T18295  
R:Feng, L.; Seymour, A.B.; Jiang, S.Y.; To, A.; Peden, A.A.; Novak, E.K.; Zhen, L.; R  
Hum. Mol. Genet. 8, 323-330, 1999  
A:Title: The beta3a subunit gene (Ap3b1) of the AP-3 adaptor complex is altered in th  
A:Reference number: 218864; MUID:99135912  
A:Accession: T18295  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1105 <FEN>  
A:Cross-references: EMBL:AF103809; NID:g3885987; PID:g3885988; PIDN:AAC78338.1  
A:Experimental source: strain CH/HeJ  
C:Genetics:  
A:Gene: Ap3b1  
A:Map position: 13

Query Match 66.7%; Score 6; DB 2; Length 1105;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVA 6  
|||||  
DB 838 VSTPVA 843

RESULT 19  
E75264  
hypothetical protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Delnoccocus radiodurans*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C:Accession: E75264  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zaleski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 A:Title: Genome sequence of the radioresistant bacterium *Delnoccocus radiodurans* R1.  
 A:Reference number: A75250; MWID:20036896  
 A:Accession: E75264  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-75 <MWID>  
 A:Cross-references: GB:AE002081; GB:AE000513; NID:96460337; PIDN:AAFI2064.1; PID:9646034  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR2520  
 A:Map position: 1

Query Match 55.6%; Score 5; DB 2; Length 75;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 VAPRO 9  
 |||||  
 DB 24 VAPRO 28

RESULT 20  
 T29802  
 hypothetical protein C06E4.2 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T29802  
 R:Du, Z.; Gattung, S.  
 A:Submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of C. elegans cosmid C06E4.  
 A:Reference number: Z20688  
 A:Accession: T29802  
 A>Status: preliminary; translated from GB/EMBL/DDBT  
 A:Molecule type: DNA  
 A:Residues: 1-82 <DDU>  
 A:Cross-references: EMBL:U41277; PIDN:AAA82477.1; CESP:C06E4.2  
 C:Genetics:  
 A:Gene: CESP:C06E4.2  
 A:Introns: 31/1

Query Match 55.6%; Score 5; DB 2; Length 82;  
 Best Local Similarity 100.0%; Pred. No. 70;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 VAPRO 9  
 |||||  
 DB 40 VAPRO 44

RESULT 21  
 T34429  
 hypothetical protein F55C7.2 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T34429  
 R:Du, Z.; Le, T.  
 A:Submitted to the EMBL Data Library, November 1996  
 A:Description: The sequence of C. elegans cosmid F55C7.  
 A:Reference number: Z21524  
 A:Accession: T34429  
 A>Status: preliminary; translated from GB/EMBL/DDBT  
 A:Molecule type: DNA  
 A:Residues: 1-88 <DDU>  
 A:Cross-references: EMBL:U80436; PIDN:AACT1107.1; GSPDB:GN00019; CESP:F55C7.2

A:Experimental source: strain Bristol N2; clone F55C7  
 C:Genetics:  
 A:Gene: CESP:F55C7.2  
 A:Map position: 1  
 A:Introns: 18/3

Query Match 55.6%; Score 5; DB 2; Length 88;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
 |||||  
 DB 31 STPVA 35

RESULT 22  
 K4RB  
 Ig kappa-B4 chain C region - rabbit  
 C:Species: *Oryctolagus cuniculus* (domestic rabbit)  
 C:Date: 31-Jan-1981 #sequence\_revision 15-Oct-1982 #text\_change 16-Aug-1996  
 C:Accession: A93971; A93891; A92176; A02122  
 R:Emorine, L.; Dreher, K.; Kindt, T.T.; Max, E.E.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 5709-5713, 1983  
 A:Title: Rabbit immunoglobulin kappa genes: structure of a germline b4 allotype J-C 1  
 A:Reference number: A93971; MWID:83300036  
 A:Accession: A93971  
 A:Molecule type: DNA  
 A:Residues: 1-103 <EMO>  
 A:Note: The sequence was determined from the germline gene  
 R:Heidmann, O.; Aulfray, C.; Cazeneuve, P.A.; Rougeon, F.  
 Proc. Natl. Acad. Sci. U.S.A. 78, 5802-5806, 1981  
 A:Title: Nucleotide sequence of constant and 3' untranslated regions of a kappa immun  
 A:Reference number: A93891; MWID:82060334  
 A:Accession: A93891  
 A:Molecule type: mRNA  
 A:Residues: 1-103 <HEI>  
 R:Chen, K.C.S.; Kindt, T.J.; Krause, R.M.  
 J. Biol. Chem. 250, 3289-3296, 1975  
 A:Title: Primary structure of the L chain from a rabbit homogeneous antibody to strep  
 A:Reference number: A92176; MWID:75133568  
 A:Accession: A92176  
 A:Molecule type: protein  
 A:Residues: 1-57, 'D', '59-103 <CHE>  
 A:Note: this chain was obtained from antibody to the specific carbohydrate of group C  
 A:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (C  
 C:Superfamily: Immunoglobulin C region; immunoglobulin homology  
 C:Keywords: heterotrimer  
 F:19-87/Domains: immunoglobulin homology <IMW>

Query Match 55.6%; Score 5; DB 1; Length 103;  
 Best Local Similarity 100.0%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
 |||||  
 DB 2 PVAPT 6

RESULT 23  
 F53275  
 Ig kappa-1 chain C region b95 allotype - rabbit (fragment)  
 C:Species: *Oryctolagus cuniculus* (domestic rabbit)  
 C:Date: 02-May-1994 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000  
 C:Accession: F53275  
 R:Ayadi, H.; Marche, P.N.; Cazeneuve, P.A.  
 Immunogenetics 34, 201-207, 1991  
 A:Title: Evolution of the rabbit immunoglobulin kappa chain genes.  
 A:Reference number: A53275; MWID:91372866  
 A:Accession: F53275  
 A>Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-104 <AYA>  
A:Note: sequence inconsistent with nucleotide translation  
A:Note: sequence extracted from NCBI backbone (NCBIN:56069, NCBIPI:56170)  
C:Superfamily: Immunoglobulin C region; Immunoglobulin homology  
C:Keywords: heterotrimer; Immunoglobulin  
F:19-87/Domain: Immunoglobulin homology <IMM>

Query Match 55.6%; Score 5; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 PVAPT 8  
11111  
DB 2 PVAPT 6

RESULT 24  
A34567  
beta-microseminoprotein precursor - human  
M:Alternate names: beta-inhibin; prostatic secretory protein; PSP-94; seminal plasma pro  
C:Species: Homo sapiens (Man)  
C:Date: 13-Jul-1990 #sequence\_revision 13-Jul-1990 #text\_change 20-Aug-1999  
C:Accession: A34567; A26451; A29777; A30984; A60673; S16238; I52682  
R:Green, C.B.; Liu, W.Y.; Kwok, S.C.M.  
Biochem. Biophys. Res. Commun. 167, 1184-1190, 1990  
A:Title: Cloning and nucleotide sequence analysis of the human beta-microseminoprotein g  
A:Reference number: A34567; MUID:90211299  
A:Accession: A34567  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-114 <GRE>  
A:Cross-references: GB:M4376; NID:9514370; PIDN:AAA59671.1; PID:9514372  
R:Mbikay, M.; Nole, S.; Fournier, S.; Benjannet, S.; Chapelaine, P.; Paradis, G.; Dubé  
DNA 6, 23-29, 1987  
A:Title: Molecular cloning and sequence of the cDNA for a 94-amino-acid seminal plasma p  
A:Reference number: A26451; MUID:87161231  
A:Accession: A26451  
A:Molecule type: mRNA  
A:Residues: 1-114 <MBI>  
A:Cross-references: GB:M15885; NID:9338414; PIDN:AAA36635.1; PID:9338415  
R:Akiyama, K.; Yoshida, Y.; Schmidt, K.; Offner, G.D.; Troxler, R.F.; Tsuda, R.; Hara, M  
Biochim. Biophys. Acta 829, 288-294, 1985  
A:Title: The amino acid sequence of human beta-microseminoprotein.  
A:Reference number: A29777; MUID:85199574  
A:Accession: A29777  
A:Molecule type: Protein  
A:Residues: 21-58; 'PT', 61-113 <AKI>  
R:Seidah, N.G.; Arbatli, N.J.; Rochement, J.; Sheth, A.R.; Chretien, M.  
FEBS Lett. 175, 349-355, 1984  
A:Title: Complete amino acid sequence of human seminal plasma beta-inhibin. Prediction c  
A:Reference number: A30984; MUID:85004133  
A:Accession: A30984  
A:Molecule type: Protein  
A:Residues: 21-112; 'G', 114 <SEI>  
R:Weber, H.; Andersson, C.; Mune, A.; Rannevik, G.; Lindstroem, C.; Lilja, H.; Fernlund  
Am. J. Pathol. 137, 593-604, 1990  
A:Title: Beta microseminoprotein is not a prostate-specific protein.  
A:Reference number: A60673; MUID:90379237  
A:Accession: A60673  
A:Molecule type: Protein  
A:Residues: 21; 'X', 23-34 <WEI>  
A:Experimental source: gastric juice  
R:Nole, S.; Mbikay, M.; Chretien, M.  
Biochim. Biophys. Acta 1089, 247-249, 1991  
A:Title: Prostatic secretory protein PSP(94): gene organization and promoter sequence in  
A:Reference number: S16237; MUID:91274557  
A:Accession: S16238  
A:Molecule type: DNA  
A:Residues: 1-114 <NOI>  
A:Cross-references: EMBL:X57928; NID:935760; PIDN:CAA41002.1; PID:9825707  
A:Note: the authors translated the codon ACT for residue 54 as Trp

R:Liu, A.Y.; Bradner, R.C.; Vessella, R.L.  
Cancer Lett. 74, 91-99, 1993  
A:Title: Decreased expression of prostatic secretory protein PSP94 in prostate cancer  
A:Reference number: I52682; MUID:94115955  
A:Accession: I52682  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-114 <RES>  
A:Cross-references: GB:S67815; NID:9460568; PIDN:AA29732.1; PID:9460569  
C:Comment: This protein is a component of seminal plasma as well as secretory fluids  
C:Genetics:  
A:Gene: GDB:MSMB  
A:Cross-references: GDB:128042; OMIM:157145  
A:Map position: 10q11.2-10q11.2  
A:Introns: 1/3; 37/1; 72/2  
C:Superfamily: seminal plasma protein  
C:Keywords: semen  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-114/Product: seminal plasma protein #status experimental <MAT>

Query Match 55.6%; Score 5; DB 2; Length 114;  
Best Local Similarity 100.0%; Pred. No. 93;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VSTPV 5  
11111  
DB 73 VSTPV 77

RESULT 25  
B49094  
methylmalonyl-CoA decarboxylase (EC 4.1.1.41) delta chain - Veillonella parvula  
C:Species: Veillonella parvula  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 05-Nov-1999  
C:Accession: B49094  
R:Huder, J.B.; Dimroth, P.  
J. Biol. Chem. 268, 24564-24571, 1993  
A:Title: Sequence of the sodium ion pump methylmalonyl-CoA decarboxylase from Veillon  
A:Reference number: A49094; MUID:94043308  
A:Accession: B49094  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-115 <HND>  
A:Cross-references: GB:L22208; NID:9415592; PIDN:AAC36621.1; PID:9415594  
C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 55.6%; Score 5; DB 2; Length 115;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 TPVAP 7  
11111  
DB 60 TPVAP 64

RESULT 26  
A72293  
conserved hypothetical protein - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: A72293  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MUID:99287316  
A:Accession: A72293  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-117 <ARN>

A:Cross-references: GB:AE001770; GB:AE000512; NID:g4981658; PIDN:AAD36200.1; PID:g498167  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM1124

Query Match 55.6%; Score 5; DB 2; Length 117;  
Best Local Similarity 100.0%; Pred. No. 95;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
|||||  
DB 30 STPVA 34

RESULT 27  
G71189  
hypothetical protein PH1791 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000  
C:Accession: G71189  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi, DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A:Reference number: A71000; MUID:98344137  
A:Accession: G71189  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-123 <KAW>  
A:Cross-references: GB:AP000007; NID:g3236134; PIDN:BAA30910.1; PID:g3258227  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by Genbank  
C:Genetics:  
A:Gene: PH1791

Query Match 55.6%; Score 5; DB 2; Length 123;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
|||||  
DB 39 VSTPV 43

RESULT 28

I54454  
Lymphocyte antigen - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 07-Jun-1996 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999  
C:Accession: I54454  
R:Hiraiwa, A.; Seyfried, C.E.; Nepom, G.T.; Milner, E.C.  
Immunogenetics 29, 186-190, 1989  
A:Title: Sequence analysis of HLA class II domains: characterization of the DQ $\alpha$ 3 family  
A:Reference number: I54454; MUID:99173071  
A:Accession: I54454  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-126 <RES>  
A:Cross-references: GB:M25325; NID:g619801; PIDN:AAA59675.1; PID:g619802  
C:Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 55.6%; Score 5; DB 2; Length 126;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
|||||  
DB 26 STPVA 30

RESULT 29  
JN0727  
hypothetical 13.4K protein - Escherichia coli

C:Species: Escherichia coli  
C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 15-May-1998  
C:Accession: JN0727  
R:Paton, A.W.; Paton, J.C.; Goldwater, P.N.; Heuzenroeder, M.W.; Manning, P.A.  
Gene 129, 87-92, 1993  
A:Title: Sequence of a variant Shiga-like toxin type-I operon of Escherichia coli O11  
A:Reference number: JN0725; MUID:93328129  
A:Accession: JN0727  
A:Molecule type: DNA  
A:Residues: 1-128 <PAT>  
A:Cross-references: GB:L04539  
A:Experimental source: serotype O11:H(-)  
C:Genetics:  
A:Gene: SLT-1

Query Match 55.6%; Score 5; DB 2; Length 128;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
|||||  
DB 56 VSTPV 60

RESULT 30

T03397  
hypothetical protein - maize (fragment)  
C:Species: Zea mays (maize)  
C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 29-Oct-1999  
C:Accession: T03397  
R:Alba, M.M.; Vinti, G.; Messegue, R.; Pages, M.  
Submitted to the EMBL Data Library, April 1997  
A:Reference number: Z14927  
A:Accession: T03397  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-132 <ALB>  
A:Cross-references: EMBL:Y12762; NID:e1045993; PIDN:CAA73301.1; PID:e314311  
A:Experimental source: strain L.W64-A

Query Match 55.6%; Score 5; DB 2; Length 132;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
|||||  
DB 101 VSTPV 105

RESULT 31

B60497  
MHC class II histocompatibility antigen RT1 D-1 beta chain precursor - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 12-Aug-1996  
C:Accession: B60497  
R:Holowachuk, E.W.; Greer, M.K.  
Diabetes 38, 267-271, 1989  
A:Title: Unaltered class II histocompatibility antigens and pathogenesis of IDDM in B  
A:Reference number: A60497; MUID:99121214  
A:Accession: B60497  
A:Molecule type: mRNA  
A:Residues: 1-134 <HOL>  
A:Superfamily: class II histocompatibility antigen; immunoglobulin homology  
C:Keywords: duplication; glycoprotein; heterodimer; transmembrane protein; transplant  
F:1-27/Domain: signal sequence #status predicted <SIG>

Query Match 55.6%; Score 5; DB 2; Length 134;



Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 STPVA 6  
Db 21 STPVA 25

## RESULT 32

E72580  
hypothetical protein APE1924 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
C:Accession: E72580  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix  
A:Reference number: A72450; MUID:99310339  
A:Accession: E72580  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-141 <KAM>  
A:Cross-references: DDBJ:AP000062; NID:95105244; PIDN:BAA80930.1; PID:Q1044716; PID:9510  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE1924

Query Match 55.6%; Score 5; DB 2; Length 141;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 VAPQ 9  
Db 66 VAPQ 70

## RESULT 33

T15056  
Photosystem I chain IV - wood tobacco

N:Alternate names: photosystem I chain PSI-E  
C:Species: Nicotiana sylvestris (wood tobacco)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: T15056  
R:Kubota, T.; Yamamoto, Y.; Obokata, J.  
Plant Physiol. 108, 1297-1298, 1995  
A:Title: Cloning of a nuclear-encoded photosystem I gene, psaEb, in Nicotiana sylvestris  
A:Reference number: 218282; MUID:95357419  
A:Accession: T15056  
A:Status: preliminary; translated from GR/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-147 <KQB>  
A:Cross-references: EMBL:DA2070; NID:9575606; PIDN:BAA07667.1; PID:Q1217601  
C:Genetics:  
A:Gene: psaEb  
A:Genome: nuclear  
A:Introns: 91/3; 113/3  
C:Function:  
A:Pathway: photosynthesis  
C:Superfamily: photosystem I chain IV  
C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match 55.6%; Score 5; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
Db 88 VSTPV 92

## RESULT 34

C96568

hypothetical protein F6D8.6 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C96568  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maitl, R.; Marzla Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talloker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: C96568  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-161 <SNO>  
A:Cross-references: GB:AE005173; NID:95903065; PIDN:AAD55624.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F6D8.6  
A:Map position: 1

Query Match 55.6%; Score 5; DB 2; Length 161;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
Db 40 PVAPT 44

## RESULT 35

B84018  
hypothetical protein BH2946 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: B84018  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20263314  
A:Accession: B84018  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-162 <SNO>  
A:Cross-references: GB:AP001517; GB:BA000004; NID:910175500; PIDN:BAB06665.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH2946

Query Match 55.6%; Score 5; DB 2; Length 162;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
Db 39 VSTPV 43

## RESULT 36

S61230  
cytochrome-c-biosynthesis heme-carrier protein cyd - Rhizobium leguminosarum

C:Species: Rhizobium leguminosarum  
C:Date: 18-Sep-1997 #sequence\_revision 18-Sep-1997 #text\_change 29-Sep-1999  
C:Accession: S61230  
R:Delgado, M.J.; Yeoman, K.H.; Wu, G.; Vargas, C.; Davies, A.; Poole, R.K.; Johnston, submitted to the EMBL Data Library, July 1995

A:Description: Characterization of the cyhCCKL genes involved in cytochrome c biogenesis  
 A:Reference number: S61229  
 A:Accession: S61230  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-165 <DEL>  
 A:Cross-references: EMBL:X89726; NID:9967064; PIDN:CAA61977.1; PID:9967066  
 C:Superfamily: Escherichia coli cytochrome-c biosynthesis heme-carrier protein come  
 C:Keywords: carrier protein; chromoprotein; heme; iron; metalloprotein  
 F:124/finding site: heme (His) (covalent) #status predicted

Query Match 55.6%; Score 5; DB 2; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
 |||||  
 DB 45 TPVAP 49

RESULT 37  
 S11770  
 ImaA protein - Listeria monocytogenes  
 C:Species: Listeria monocytogenes  
 C:Date: 03-May-1994 #sequence\_revision 20-Feb-1995 #text\_change 24-Nov-1999  
 C:Accession: S11770  
 R:Goehmann, S.; Lehnlester-Maehter, M.; Schlitz, E.; Goebel, W.; Chakraborty, T.  
 M:Microbiol. 4, 1091-1099, 1990  
 A:Title: Characterization of a Listeria monocytogenes-specific protein capable of inducing  
 A:Reference number: S11769; MUID:91041717  
 A:Accession: S11770  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-170 <GOE>  
 C:Genetics:  
 A:Gene: ImaA  
 C:Superfamily: Listeria monocytogenes ImaA protein

Query Match 55.6%; Score 5; DB 2; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
 |||||  
 DB 164 TPVAP 168

RESULT 38  
 E83871  
 hypothetical protein BH1773 [imported] - Bacillus halodurans (strain C-125)  
 C:Species: Bacillus halodurans  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
 C:Accession: E83871  
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A:Reference number: A83650; MUID:20263314  
 A:Accession: E83871  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-170 <STO>  
 A:Cross-references: GB:AP001513; GB:BA000004; NID:910174345; PIDN:BA05492.1; GSPDB:GN00  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: BH1773

Query Match 55.6%; Score 5; DB 2; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTPV 5  
 |||||  
 DB 124 VSTPV 128

RESULT 39  
 T50849  
 hypothetical protein 110K5.14 [imported] - sorghum (fragment)  
 C:Species: Sorghum bicolor (sorghum)  
 C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
 C:Accession: T50849  
 R:Titkonov, A.P.; SanMiguel, P.J.; Nakajima, Y.; Gorenstein, N.M.; Bennettzen, J.L.; A  
 Proc. Natl. Acad. Sci. U.S.A. 96, 7409-7414, 1999  
 A:Title: Colinearity and its exceptions in orthologous adh regions of maize and sorgh  
 A:Reference number: Z25261; MUID:99307423  
 A:Accession: T50849  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-175 <TRK>  
 A:Cross-references: EMBL:AF124045; PIDN:AD43043.1  
 A:Experimental source: cultivar BTx623  
 C:Genetics:  
 A:Map position: J  
 A:Note: 110K5.14

Query Match 55.6%; Score 5; DB 2; Length 175;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
 |||||  
 DB 68 TPVAP 72

RESULT 40  
 S57789  
 hypothetical protein (clone ES2A) - barley  
 C:Species: Hordeum vulgare (barley)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: S57789  
 R:Speulman, E.; Salamini, F.  
 Plant Mol. Biol. 28, 915-926, 1995  
 A:Title: GA(3)-regulated cDNAs from Hordeum vulgare leaves.  
 A:Reference number: S57787; MUID:95367651  
 A:Accession: S57789  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-177 <SPE>  
 A:Cross-references: EMBL:X79466; NID:9929668; PIDN:CAA55976.1; PID:9929669  
 C:Superfamily: barley hypothetical protein (clone ES2A)

Query Match 55.6%; Score 5; DB 1; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
 |||||  
 DB 140 TPVAP 144

RESULT 41  
 KKRT  
 kappa-casein precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 22-Jun-1999  
 C:Accession: A03115  
 R:Nakhasi, H.L.; Grantam, F.H.; Guillino, P.M.  
 J. Biol. Chem. 259, 14894-14898, 1984  
 A:Title: Expression of kappa-casein in normal and neoplastic rat mammary gland is und  
 A:Reference number: A03115; MUID:85054984  
 A:Accession: A03115

A:Molecule type: mRNA  
A:Residues: 1-178 <NKA>  
C:Cross-references: GB:R02598; NID:g203320; PIDN:AAA40880.1; PID:g203321  
C:Comment: kappa casein stabilizes micelle formation, preventing casein precipitation in  
C:Comment: Pro lactase modules the production of kappa casein mRNA.  
C:Superfamily: kappa-casein  
C:Keywords: glycoprotein; mammary gland; phosphoprotein  
F:1-21/Domains: signal sequence #status predicted <SIG>  
F:112-178/Product: kappa-casein #status predicted <KCA>  
F:113/Binding site: carbohydrate (asn) (covalent) #status predicted  
F:115-116/Cleavage site: Phe-Leu (chymosin) #status predicted  
F:120,145,165/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 55.6%; Score 5; DB 1; Length 178;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTPV 5  
|||||  
DB 139 VSTPV 143

RESULT 42  
S56300  
hypothetical protein YFR045w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein R014  
C:Species: Saccharomyces cerevisiae  
C:Date: 02-Sep-1995 #sequence\_revision 19-Oct-1995 #text\_change 20-Jun-2000  
C:Accession: S56300; S62256; S63792  
R:Murakami, Y.; Naitou, M.; Hagihara, H.; Shibata, T.; Ozawa, M.; Sasamura, S.I.; Sasani  
submitted to the EMBL Data Library, May 1995  
A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces ce  
A:Reference number: S56186  
A:Accession: S56300  
A:Molecule type: DNA  
A:Residues: 1-178 <MUR>  
A:Cross-references: EMBL:D50617; NID:g836685; PIDN:BAAO9284.1; PID:g836800; MIPS:YFR045w  
R:Murakami, Y.  
submitted to the EMBL Data Library, December 1994  
A:Reference number: S62230  
A:Accession: S62230  
A:Molecule type: DNA  
A:Residues: 1-178 <MUR>  
A:Cross-references: EMBL:D44597; NID:g871938; PIDN:BAAO8011.1; PID:g871944  
R:EKI, T.; Naitou, M.; Hagihara, H.; Ozawa, M.; Sasamura, S.I.; Sasamura, M.; Tsuchiya,  
Teast 12, 149-167, 1996  
A:Title: Analysis of a 36.2 kb DNA sequence including the right telomere of chromosome V  
A:Reference number: S63787; MUID:96287652  
A:Accession: S63792  
A:Status: nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 101-154 <EKI>  
A:Cross-references: EMBL:D44597  
C:Genetics:  
A:Map position: 6R  
A:Note: YFR045w  
C:Superfamily: hypothetical protein YFR045w; ADP-ATP carrier protein repeat homology  
C:Keywords: duplication; transmembrane protein

Query Match 55.6%; Score 5; DB 2; Length 178;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 STPVA 6  
|||||  
DB 27 STPVA 31

RESULT 43  
S01400  
H+-transporting ATP synthase (EC 3.6.1.34) delta chain - thermophilic bacterium PS-3

C:Species: thermophilic bacterium PS-3  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 22-Jun-1999  
C:Accession: S01400  
R:Ohba, S.; Yoda, M.; Ishizuka, M.; Hirata, H.; Hamamoto, T.; Otawara-Hamamoto, Y.;  
Biochim. Biophys. Acta 933, 141-155, 1988  
A:Title: Sequence and over-expression of subunits of adenosine triphosphate synthase  
A:Reference number: S01397; MUID:88163679  
A:Accession: S01400  
A:Molecule type: DNA  
A:Residues: 1-179 <OHT>  
A:Cross-references: EMBL:X07804; NID:g45808; PIDN:CAA30651.1; PID:g45813  
A:Note: part of this sequence, including the amino end of the mature protein, was con  
C:Superfamily: H+-transporting ATP synthase delta chain  
C:Keywords: ATP biosynthesis; hydrolase  
F:1-179/Product: H+-transporting ATP synthase delta chain #status experimental <MAT>

Query Match 55.6%; Score 5; DB 2; Length 179;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTPV 5  
|||||  
DB 69 VSTPV 73

RESULT 44  
G84789  
hypothetical protein At2g37200 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: G84789  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Beilto, M.I.; Town, C.D.; Fujii, C.Y  
M.; Koo, H.; Moffat, K.S.; Conlin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: G84789  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-179 <STO>  
A:Cross-references: GB:AE002093; NID:g4371281; PIDN:AA018139.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g37200  
A:Map position: 2

Query Match 55.6%; Score 5; DB 2; Length 179;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PVAPT 8  
|||||  
DB 15 PVAPT 19

RESULT 45  
A05080  
kappa-casein precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 23-Feb-1997  
C:Accession: A05080  
R:Thompson, M.D.; Dave, J.R.; Nakhasi, H.L.  
DNA 4, 263-271, 1985  
A:Reference number: A05080; MUID:86004062  
A:Accession: A05080  
A:Molecule type: mRNA  
A:Residues: 1-181 <THO>  
A:Note: the authors translated the codon GAU for residue 27 as Asn  
C:Superfamily: kappa-casein  
C:Keywords: phosphoprotein

Query Match 55.6%; Score 5; DB 2; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 |||||  
 Db 142 VSTPV 146

## RESULT 46

D96572  
 Protein F12M16.6 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: D96572  
 R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Li, J.H.; Hughes, B.; Hulzar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: D96572  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-185 <STO>  
 A:Cross-references: GB:AE005173; NID:97769876; PIDN:AAF69554.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: F12M16.6  
 A:Map position: 1

Query Match 55.6%; Score 5; DB 2; Length 185;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
 |||||  
 Db 12 PVAPT 16

## RESULT 47

H86404  
 Probable lipid transfer protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: H86404  
 R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Li, J.H.; Hughes, B.; Hulzar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719  
 A:Accession: H86404  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-193 <STO>  
 A:Cross-references: GB:AE005172; NID:911024871; PIDN:AAG26955.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 55.6%; Score 5; DB 2; Length 193;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
 |||||  
 Db 139 TPVAP 143

## RESULT 48

D85875  
 Probable minor fibrial subunit Z3598 [imported] - Escherichia coli (strain 0157:H7)  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
 C:Accession: D85875  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May, J.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda, Nature 409, 523-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: D85875  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-197 <STO>  
 A:Cross-references: GB:AE005174; NID:912516700; PIDN:AAG57464.1; GSPDB:GN00145; UWGP:  
 A:Experimental source: strain 0157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: Z3598

Query Match 55.6%; Score 5; DB 2; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 |||||  
 Db 48 VSTPV 52

## RESULT 49

F86832  
 Hypothetical protein yrbB [imported] - Lactococcus lactis subsp. lactis (strain IL140  
 C:Species: Lactococcus lactis subsp. lactis  
 C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 23-Mar-2001  
 C:Accession: F86832  
 R:Boletín, A.; Winkler, P.; Manger, S.; Jallón, O.; Malarme, K.; Weissenbach, J.; Eh Genome Res. in press, 2001  
 A:Title: The complete genome sequence of the lactic acid bacterium.  
 A:Reference number: A86625  
 A:Accession: F86832  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-199 <STO>  
 A:Cross-references: GB:AE005176; NID:912724674; PIDN:AAK05760.1; GSPDB:GN00146  
 A:Experimental source: strain IL1403  
 C:Genetics:  
 A:Gene: yrbB

Query Match 55.6%; Score 5; DB 2; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 VAPRO 9  
 |||||  
 Db 96 VAPRO 100

## RESULT 50

S69748  
 Hypothetical protein YDR537c - yeast (Saccharomyces cerevisiae)  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 23-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 04-Mar-2000  
 C:Accession: S69748

R:Dielectric, F.S.; Mulligan, J.; Allen, E.; Araujo, R.; Aviles, E.; Berno, A.; Carpenter, H.; Lin, D.; Mosedale, D.; Nakahara, K.; Namath, A.; Oefner, P.; Oh, C.; Petel, F.X.; Ro

submitted to the EMBL Data Library, December 1995

A:Reference number: S62017

A:Accession: S69748

A:Molecule type: DNA

A:Residues: 1-201 <DIE>

A:Cross-references: EMBL:U43834; NID:g1165292; PID:g2194177; GSPDB:GN00004; MIPS:YDR537C

C:Genetics:

A:Gene: MIPS:YDR537C

A:Map position: 4R

C:Superfamily: Saccharomyces cerevisiae hypothetical protein YDR537C

Query Match 55.6%; Score 5; DB 2; Length 201;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7

DB 121 TPVAP 125

Search completed: August 15, 2001, 12:34:31  
Job time: 96 sec

***This Page Blank (uspio)***

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:34:50 ; Search time 20.38 seconds  
(without alignments)  
15.128 Million cell updates/sec

Title: US-09-372-036-26

Perfect score: 9

Sequence: 1 VSTPVAProQ 9

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size: 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database: SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	100.0	484	1 P60_LISMO	P21171 listeria mo
2	6	66.7	155	1 PRS1_SOLITU	P17641 solanum tub
3	6	66.7	187	1 VGG_BPPIK	Q38042 bacterioph
4	6	66.7	239	1 SOD3_PLEBO	P50060 plectonema
5	6	66.7	453	1 GAT4_HELPJ	Q92113 helicobacte
6	6	66.7	453	1 GAT4_HELPJ	P56114 helicobacte
7	6	66.7	711	1 CADA_LISMO	O60048 listeria mo
8	6	66.7	865	1 CPN_DROME	O02910 drosophila
9	6	66.7	1540	1 RML_DROME	O9y7h4 drosophila
10	5	55.6	103	1 KAC4_RABIT	P01840 oryctolagus
11	5	55.6	104	1 ST12_MOUSE	O62401 mus musculu
12	5	55.6	114	1 PSSP_HUMAN	P08118 homo sapien
13	5	55.6	166	1 YELJ_SCHPO	O13883 schizosacch
14	5	55.6	178	1 CASK_RAT	P04468 rattus norv
15	5	55.6	178	1 YFL5_YEAST	P43617 saccharomyc
16	5	55.6	179	1 ARPD_BAC3	P09220 bacillus ps
17	5	55.6	181	1 CASK_MOUSE	P06796 mus musculu
18	5	55.6	217	1 V33P_ADE41	P19416 human adeno
19	5	55.6	218	1 RPE_TREPA	O66107 treponema p
20	5	55.6	225	1 Y574_AERPE	O9yeka aeropyrum p
21	5	55.6	229	1 PRL_BOVIN	P01239 bos taurus
22	5	55.6	229	1 PRL_CAPRI	O28318 capra hircu
23	5	55.6	229	1 PRL_SHEEP	P01240 ovis aries
24	5	55.6	261	1 HB21_HUMAN	P01918 homo sapien
25	5	55.6	261	1 HB24_HUMAN	P01920 homo sapien
26	5	55.6	263	1 GP3D_CHLTR	P10557 chlamydia t
27	5	55.6	264	1 PSBP_FRING	O49080 fillifaria
28	5	55.6	268	1 HB2X_HUMAN	P05538 homo sapien
29	5	55.6	271	1 YSM4_CAEEL	O10124 caenorhabdi
30	5	55.6	277	1 Y4OR_RHISN	P35603 rhizobium s
31	5	55.6	303	1 PPV_DROME	O27884 drosophila
32	5	55.6	305	1 PPP6_HUMAN	O00743 homo sapien
33	5	55.6	305	1 PPP6_RAT	O64620 rattus norv

34	5	55.6	309	1 POLG_HCVH7	P27955 hepatitis c
35	5	55.6	314	1 ACCO_MALDO	O00985 malus domes
36	5	55.6	321	1 SPLR_NPYOP	O65328 oryza psen
37	5	55.6	323	1 PF27_MOUSE	P52875 mus musculu
38	5	55.6	325	1 HMGL_HUMAN	P35914 homo sapien
39	5	55.6	325	1 MODD_MYCBO	O30620 mycobacteri
40	5	55.6	325	1 MODD_MYCBO	O50906 mycobacteri
41	5	55.6	336	1 YB56_XENLA	P21574 xenopus lae
42	5	55.6	337	1 RL3_METHA	O26110 methanobact
43	5	55.6	354	1 YWA3_MYCBO	O02279 mycobacteri
44	5	55.6	356	1 HIS8_ECOLI	P06986 escherichia
45	5	55.6	359	1 HIS8_SALTY	P10369 salmonella
46	5	55.6	362	1 SRK2_SPOA	P42688 spongilla l
47	5	55.6	366	1 RRPO_REOVD	P03526 reovirus (t
48	5	55.6	366	1 RRPO_REOVD	P12002 reovirus (t
49	5	55.6	366	1 RRPO_REOVL	P07940 reovirus (t
50	5	55.6	366	1 YW85_YEAST	O04806 saccharomyc
51	5	55.6	367	1 MTG1_CHV13	P36216 chlorella v
52	5	55.6	367	1 NUIJ_PODAN	P19041 podopora a
53	5	55.6	377	1 TTL_BOVIN	P38584 bos taurus
54	5	55.6	379	1 TTL_PIG	P38160 sus scrofa
55	5	55.6	380	1 SUC2_ARCFU	O28097 archaeoglob
56	5	55.6	380	1 YMH7_CAEEL	P34474 caenorhabdi
57	5	55.6	386	1 YG5V_YEAST	P53334 saccharomyc
58	5	55.6	390	1 EMRA_ECOLI	P27303 escherichia
59	5	55.6	391	1 THAB_PAROL	O91242 paratichthy
60	5	55.6	394	1 ALR2_BACSU	P94494 bacillus su
61	5	55.6	414	1 HRTD_CROAT	P15167 croctalus at
62	5	55.6	423	1 ODB2_CAEEL	O18964 caenorhabdi
63	5	55.6	423	1 ODB2_PSEPU	P09062 pseudomonas
64	5	55.6	428	1 SCK_HUMAN	P98077 homo sapien
65	5	55.6	432	1 BRAC_XENLA	P24781 xenopus lae
66	5	55.6	433	1 RAS2_RLUDA	P41768 kluyveromyc
67	5	55.6	442	1 TTG_CHLYR	O84713 chlamydia t
68	5	55.6	453	1 YOTA_CAEEL	P34656 caenorhabdi
69	5	55.6	458	1 RRA_XENLA	P51126 xenopus lae
70	5	55.6	468	1 PCOL_MOUSE	O61398 mus musculu
71	5	55.6	468	1 PCOL_RAT	O08628 rattus norv
72	5	55.6	470	1 GUNA_FREDI	P33035 tremella d
73	5	55.6	473	1 GUNA_SYNT3	P77961 synechocyst
74	5	55.6	473	1 YCLG_BACSU	P94405 bacillus su
75	5	55.6	474	1 YDCC_STRD7	O94697 streptomyc
76	5	55.6	474	1 VP61_NPYOP	O10270 oryza psen
77	5	55.6	478	1 HRTF_CROAT	P34182 croctalus at
78	5	55.6	480	1 DISA_TRIGA	P15503 trimeresuru
79	5	55.6	488	1 EXON_HSV62	P24447 human herpe
80	5	55.6	488	1 EXON_HSV62	P52448 human herpe
81	5	55.6	488	1 PE24_RABIT	O28691 oryctolagus
82	5	55.6	488	1 SPL1_CANAL	P87185 candida alb
83	5	55.6	498	1 ACHP_HUMAN	P30926 homo sapien
84	5	55.6	501	1 YH92_CAEEL	O23236 caenorhabdi
85	5	55.6	504	1 ARF5_HUMAN	O13568 homo sapien
86	5	55.6	509	1 CPT7_BOVIN	P05185 bos taurus
87	5	55.6	509	1 CX56_CHICK	P24415 gallus galli
88	5	55.6	509	1 PSBB_GUTRH	O78511 gallinaria
89	5	55.6	509	1 VLI_HPV49	P38742 human papil
90	5	55.6	510	1 PUR1_YEAST	P04046 saccharomyc
91	5	55.6	512	1 VLI_HPV24	P05090 human papil
92	5	55.6	514	1 VLI_HPV08	P06417 human papil
93	5	55.6	514	1 VLI_HPV47	P22424 human papil
94	5	55.6	515	1 VLI_HPV50	P50818 human papil
95	5	55.6	516	1 VLI_HPV05	P06917 human papil
96	5	55.6	516	1 VLI_HPV20	P50762 human papil
97	5	55.6	516	1 VLI_HPV30	P50812 human papil
98	5	55.6	517	1 VLI_HPV12	P36733 human papil
99	5	55.6	517	1 VLI_HPV25	O02051 human papil
100	5	55.6	518	1 VLI_HPV14	P36734 human papil
101	5	55.6	518	1 VLI_HPV21	P50787 human papil
102	5	55.6	525	1 VLI_HPV58	P26537 human papil
103	5	55.6	546	1 VLI_HPV19	O02050 human papil
104	5	55.6	552	1 COX1_SYNTU	O02676 synechococc
105	5	55.6	556	1 CAG1_DROME	O94360 drosophila
106	5	55.6	562	1 CHS5_CANAL	O74161 candida alb

107	5	55.6	601	1	DNM_MYCGE	P47442 mycoplasma	180	4	44.4	48	1	TXA2_RADPA	P01534 radianthus
108	5	55.6	604	1	BAP3_YEAST	P41815 saccharomyc	181	4	44.4	60	1	YTR2_SPIAU	P22042 spirochaeta
109	5	55.6	609	1	BAP3_YEAST	P38084 saccharomyc	182	4	44.4	62	1	UCRX_BOVIN	P00130 bos taurus
110	5	55.6	612	1	SYT_HELPJ	Q92903 helicobacte	183	4	44.4	63	1	ANP2_MACAM	P19609 macrozoarc
111	5	55.6	612	1	SYT_HELPJ	P56071 helicobacte	184	4	44.4	64	1	ANP2_MACAM	P19608 macrozoarc
112	5	55.6	614	1	YEOH_ECOLI	P33353 escherichia	185	4	44.4	65	1	YNOL_RHIFR	P33214 thizobium f
113	5	55.6	634	1	VEI_HPVS1	P26544 human papil	186	4	44.4	66	1	YVGR_STRCO	Q05954 streptomyce
114	5	55.6	634	1	2147_MOUSE	O61510 mus musculi	187	4	44.4	66	1	YYCD_BACSV	P37480 bacillus su
115	5	55.6	642	1	DP3X_SALTY	P74876 salmonella	188	4	44.4	67	1	HJAI_METJA	Q57632 methanococc
116	5	55.6	643	1	DP3X_ECOLI	P66710 escherichia	189	4	44.4	67	1	HJAI_METJA	Q58342 methanococc
117	5	55.6	644	1	ARNT_DROME	O15945 drosophilla	190	4	44.4	67	1	HJAI_METJA	Q58655 methanococc
118	5	55.6	648	1	YDEE_SCHPO	O10447 schizosacch	191	4	44.4	67	1	HJAI_METJA	Q60264 methanococc
119	5	55.6	649	1	DREB_HUMAN	O16643 homo sapien	192	4	44.4	70	1	F261_MOUSE	P70266 mus musculi
120	5	55.6	671	1	ALYS_ENTFA	P37710 enterococcu	193	4	44.4	71	1	RL30_MYCLE	O33001 mycobacteri
121	5	55.6	707	1	BMP1_XENTIA	P98070 xenopus lae	194	4	44.4	72	1	CCAA_ECOLI	Q46959 escherichia
122	5	55.6	760	1	AHM4_ARATH	Q9szw5 arabidopsis	195	4	44.4	72	1	VG35_BMML5	O05235 mycobacteri
123	5	55.6	767	1	ANYT_YEAST	P53852 saccharomyc	196	4	44.4	77	1	RUXG_YEAST	P40204 saccharomyc
124	5	55.6	781	1	YGS5_YEAST	P53316 saccharomyc	197	4	44.4	79	1	PSBH_CHDVU	P56323 chlorella v
125	5	55.6	790	1	SMY2_YEAST	P32909 saccharomyc	198	4	44.4	80	1	GONI_CLAGA	P33439 clarifas gar
126	5	55.6	794	1	PMS1_SCHPO	P54280 schizosacch	199	4	44.4	87	1	ANP1_MACAM	P07457 macrozoarc
127	5	55.6	806	1	NUOG_MYCTU	O05192 mycobacteri	200	4	44.4	87	1	ANP5_MACAM	P19607 macrozoarc
128	5	55.6	808	1	FTFB_DROME	P95175 mycobacteri	201	4	44.4	87	1	ANP5_MACAM	P19604 macrozoarc
129	5	55.6	811	1	Y104_SYNY3	P54371 synchocyst	202	4	44.4	93	1	IM05_ARATH	Q09x99 arabidopsis
130	5	55.6	817	1	VRP1_YEAST	P37370 saccharomyc	203	4	44.4	93	1	YQIC_CABEL	Q09283 caenorhabdi
131	5	55.6	834	1	POLG_ENNGO	P12296 mengo encp	204	4	44.4	98	1	P531_MOUSE	P70399 mus musculi
132	5	55.6	835	1	INVA_YEREN	P19196 yeislinia en	205	4	44.4	99	1	SY08_BOVIN	Q09141 bos taurus
133	5	55.6	836	1	Y167_HUMAN	O99490 homo sapien	206	4	44.4	100	1	POL_SIV33	P12501 simlan immu
134	5	55.6	845	1	RRPO_IPNVJ	P22173 infectious	207	4	44.4	101	1	VNST_BUNYW	P16494 bunyamwera
135	5	55.6	866	1	DYN2_MOUSE	P39054 mus musculi	208	4	44.4	101	1	VNST_MAGV	P16605 maguari vlr
136	5	55.6	870	1	DYN2_HUMAN	P50570 homo sapien	209	4	44.4	102	1	SGP3_CHRVI	O52055 chromatium
137	5	55.6	870	1	DYN2_RAT	P39052 rattus norv	210	4	44.4	102	1	VE7_PAPVE	P01331 deer papill
138	5	55.6	901	1	POLG_ENNG3	P32540 mengo encp	211	4	44.4	102	1	VE7_PAPVE	P11372 european el
139	5	55.6	909	1	CTIA_FUSO	P52958 fusarium so	212	4	44.4	103	1	LAC_CHICK	P20763 gallus gall
140	5	55.6	930	1	PRGR_RABIT	P06186 oryctolagus	213	4	44.4	103	1	LAC_BOVIN	P01352 bos taurus
141	5	55.6	933	1	PRGR_HUMAN	P06401 homo sapien	214	4	44.4	104	1	GAST_SHEEP	O02686 ovis aries
142	5	55.6	939	1	ST20_YEAST	O03487 saccharomyc	215	4	44.4	104	1	YR7H_ECOLI	P21322 escherichia
143	5	55.6	951	1	AHM5_ARATH	Q9szw4 arabidopsis	216	4	44.4	105	1	LAC_HUMAN	P01842 homo sapien
144	5	55.6	954	1	Y442_HUMAN	Q9upv9 homo sapien	217	4	44.4	105	1	NDQ1_BRAJA	P26024 bradyrhizob
145	5	55.6	983	1	BIR1_YEAST	P47134 saccharomyc	218	4	44.4	105	1	YMO3_CABEL	P34494 caenorhabdi
146	5	55.6	983	1	Y144_HUMAN	O14157 homo sapien	219	4	44.4	106	1	KACH_RABIT	P01839 oryctolagus
147	5	55.6	986	1	BMP1_HUMAN	P13487 homo sapien	220	4	44.4	106	1	TBCA_YEAST	P48606 saccharomyc
148	5	55.6	991	1	BMP1_MOUSE	P98063 mus musculi	221	4	44.4	107	1	U139_ARATH	Q9s888 arabidopsis
149	5	55.6	1009	1	DPOL_HSVSA	P24907 herpesvltus	222	4	44.4	108	1	HDEB_ECOLI	P26659 escherichia
150	5	55.6	1018	1	HMW1_MYCPN	O50365 mycoplasma	223	4	44.4	109	1	CORT_MOUSE	P56465 mus musculi
151	5	55.6	1027	1	POL_SIVC2	P17283 chimpanzee	224	4	44.4	112	1	CL34_DROMI	Q01774 drosophila
152	5	55.6	1057	1	SEFI_YEAST	P34228 saccharomyc	225	4	44.4	113	1	Y12K_BPPHH	P23788 bacteriophla
153	5	55.6	1141	1	MYFS_HUMAN	O00872 homo sapien	226	4	44.4	114	1	CU05_HOMAM	P81389 homarus ame
154	5	55.6	1171	1	NIEF_KLEPN	P03833 klebsiella	227	4	44.4	114	1	ET3_RABIT	P19998 oryctolagus
155	5	55.6	1172	1	AHM2_ARATH	O64474 arabidopsis	228	4	44.4	115	1	GUAN_RAT	P28902 rattus norv
156	5	55.6	1178	1	RPOB_MYCTU	P47766 mycobacteri	229	4	44.4	115	1	KV51_MOUSE	P01642 mus musculi
157	5	55.6	1179	1	RPOB_MYCLE	P30760 mycobacteri	230	4	44.4	116	1	B2MG_BRARE	Q04475 brachydanio
158	5	55.6	1213	1	T2D2_DROME	O24335 drosophila	231	4	44.4	116	1	GUAN_MOUSE	P33680 mus musculi
159	5	55.6	1253	1	POLS_SFV	P03315 semliki for	232	4	44.4	117	1	NLT5_ORYSA	O65091 oryza sativ
160	5	55.6	1293	1	IRR_HUMAN	P14616 homo sapien	233	4	44.4	119	1	MLP3_BOVIN	O4315 bos taurus
161	5	55.6	1300	1	IRR_CAYPO	P14617 cavia porce	234	4	44.4	120	1	CU19_ARADI	P80515 araneus dia
162	5	55.6	1337	1	P152_YEAST	P39685 saccharomyc	235	4	44.4	120	1	YRM7_CABEL	Q09416 caenorhabdi
163	5	55.6	1500	1	SSP5_STRGN	P16952 streptococ	236	4	44.4	121	1	RLA2_TAESO	P50879 taenia soli
164	5	55.6	1627	1	ADP1_MYCPN	P11311 mycoplasma	237	4	44.4	121	1	YHIF_AOUAE	O66336 aquifex aeo
165	5	55.6	1670	1	CA34_HUMAN	O01955 homo sapien	238	4	44.4	124	1	SGP2_CHRVI	O52179 chromatium
166	5	55.6	1736	1	ZOL_HUMAN	O07157 homo sapien	239	4	44.4	124	1	VNS4_CVMS	P29075 murine coro
167	5	55.6	1781	1	AKAC_HUMAN	O02992 homo sapien	240	4	44.4	125	1	FLIO_SALTY	P54699 salmonella
168	5	55.6	1794	1	YAVI_SCHPO	O10172 schizosacch	241	4	44.4	126	1	CU24_ARADI	P80516 araneus dia
169	5	55.6	1833	1	ZEP2_HUMAN	P31629 homo sapien	242	4	44.4	127	1	CB21_EUGGR	P12317 euglena gra
170	5	55.6	2208	1	POLN_MANCV	O69014 manchester	243	4	44.4	127	1	CU26_ARADI	P480517 araneus dia
171	5	55.6	2290	1	POLG_EMKV	P03304 encephalom	244	4	44.4	127	1	YCBK_HAEIN	P44384 haemophilus
172	5	55.6	2292	1	POLG_EMKVB	P17593 encephalom	245	4	44.4	127	1	YFID_HAEIN	P44455 haemophilus
173	5	55.6	2422	1	POLG_EMKVD	P17594 encephalom	246	4	44.4	129	1	IGF2_MUSVI	P41694 musceta vis
174	5	55.6	2440	1	NCRI_HUMAN	O75376 homo sapien	247	4	44.4	129	1	RL7_TREPA	O83368 treponema p
175	5	55.6	2476	1	ZAN_PIG	O28963 sus scrofa	248	4	44.4	130	1	RS8_METVA	P14038 methanococ
176	5	55.6	3133	1	HMCT_BOMO	P98092 bombyx mori	249	4	44.4	130	1	Y056_MYCTU	P71056 mycobacteri
177	5	55.6	3418	1	BRG2_HUMAN	P51587 homo sapien	250	4	44.4	131	1	YN09_YEAST	P53841 saccharomyc
178	4	44.4	48	1	TXA1_RADMA	P30831 radianthus	251	4	44.4	132	1	HEX9_ADE40	P48312 human adeno
179	4	44.4	48	1	TXA2_RADMA	P30783 radianthus	252	4	44.4	132	1	IF5A_METJA	O58625 methanococ



253	4	44.4	133	1	HEX9_ADE41	P32539	human	adeno	326	4	44.4	164	1	YOJ3_CAEEL	P34626	caenorhabdi
254	4	44.4	133	1	PROF_HELAN	081982	helianthus		327	4	44.4	167	1	VNS3_BMDNV	P05841	bombyx dens
255	4	44.4	133	1	RS8_LEPIN	094482	leptosira		328	4	44.4	168	1	HEM2_SALTY	P37408	salmonella
256	4	44.4	133	1	RS8_METTH	026126	methanobact		329	4	44.4	168	1	YNP7_YEAST	P33897	saccharomyc
257	4	44.4	135	1	PTLG_PSEAE	P46384	pseudomonas		330	4	44.4	171	1	IF3_LISMO	O53084	listeria mo
258	4	44.4	136	1	EXB2_XANCP	034258	xanthomonas		331	4	44.4	171	1	YSYM_RHIET	O08815	rhizobium e
259	4	44.4	137	1	NIFB_KLEDOX	P56265	klebsiella		332	4	44.4	172	1	RUB2_PSEOL	P00272	pseudomonas
260	4	44.4	138	1	CXA4_PIG	029559	sus scrofa		333	4	44.4	174	1	NEUJ_RAT	P12760	rattus norv
261	4	44.4	138	1	NEDR_MICVI	002835	micromonos		334	4	44.4	175	1	LPFE_SALTY	P43664	salmonella
262	4	44.4	140	1	MRP_CLOPE	P53381	clostridium		335	4	44.4	175	1	RBS_AEGSO	Q38793	aegilops sq
263	4	44.4	140	1	Y14K_CSMY	P18920	chlortis str		336	4	44.4	176	1	CON8_NEUCR	P10166	neurospora
264	4	44.4	140	1	YB60_HELPU	025777	helicobacte		337	4	44.4	176	1	IPYR_AERPE	O9Y645	aeropyrum p
265	4	44.4	141	1	AT92_HUMAN	006055	homo sapien		338	4	44.4	178	1	ATPD_BACCA	P41011	baecillus ca
266	4	44.4	141	1	MPL3_RAT	062625	rattus norv		339	4	44.4	179	1	REK27_TOBAC	P01055	nicotiana t
267	4	44.4	142	1	IL3_CALJA	Q28334	callithrix		340	4	44.4	180	1	IGF2_TOBAC	P01344	homo sapien
268	4	44.4	142	1	IL3_SAGOE	P51445	saguinus oe		341	4	44.4	180	1	YMER_STYAU	P08655	staphylococ
269	4	44.4	142	1	RECX_PSEPU	P37862	pseudomonas		342	4	44.4	180	1	YPKW_THREU	P49694	thermonomos
270	4	44.4	142	1	Y83A_METJA	P81322	methanococc		343	4	44.4	181	1	IGF2_HORSE	P51459	equus cabal
271	4	44.4	142	1	Y387_MYCTU	010870	mycobacteri		344	4	44.4	181	1	IGF2_PIG	P23695	sus scrofa
272	4	44.4	143	1	AT92_BOVIN	P07926	bos taurus		345	4	44.4	181	1	NUSG_BUCAI	P57151	buchnera ap
273	4	44.4	143	1	AT92_SHEEP	006056	ovis aries		346	4	44.4	182	1	RBS_MANES	Q42913	manihot esc
274	4	44.4	143	1	IL4_MERUN	P47966	meriones un		347	4	44.4	183	1	OLEC_BRANA	P29526	brassica na
275	4	44.4	143	1	PER_DROPI	Q25206	drosophila		348	4	44.4	184	1	NUSG_BORBU	O51355	borrelia bu
276	4	44.4	143	1	SGCA_ECOLI	P39363	escherichia		349	4	44.4	185	1	SC6_SCHCO	O74300	schlicophyll
277	4	44.4	144	1	HEX9_ADEI2	P03284	human	adeno	350	4	44.4	186	1	ARDH_LEIDO	O05885	leishmania
278	4	44.4	144	1	RISB_PHOLE	001994	photobacter		351	4	44.4	187	1	BNZB_PSEPU	P08085	pseudomonas
279	4	44.4	145	1	YNR9_YEAST	P53880	saccharomyc		352	4	44.4	187	1	MAUE_METEX	Q49123	methylobact
280	4	44.4	146	1	H2B_TOBAC	P93354	nicotiana t		353	4	44.4	187	1	TOD2_PSEPU	P13451	pseudomonas
281	4	44.4	146	1	HBB1_ONCMY	P02142	oncorhynchu		354	4	44.4	190	1	WS3_HUMAN	O00399	homo sapien
282	4	44.4	147	1	YACA_ECOLI	P10409	escherichia		355	4	44.4	190	1	YBAY_ECOLI	P42177	escherichia
283	4	44.4	151	1	SOD1_LYCES	P14830	lycopersico		356	4	44.4	190	1	YELM_SALTY	P40709	salmonella
284	4	44.4	152	1	CORA_RAT	O65332	rattus norv		357	4	44.4	191	1	HGRE_RHILV	P27650	rhizobium l
285	4	44.4	152	1	PSAL_SYNP2	Q54755	synechococc		358	4	44.4	191	1	PGHD_RHILV	Q34865	canis famli
286	4	44.4	152	1	PYR1_PYRHO	O58452	pyrococcus		359	4	44.4	191	1	RL3_HELPJ	Q5474	helicobacte
287	4	44.4	153	1	PSAL_PORPU	P51222	porphyra pu		360	4	44.4	191	1	RL3_HELPJ	P56031	helicobacte
288	4	44.4	153	1	RECX_PSEAE	P37860	pseudomonas		361	4	44.4	192	1	KADA_METVO	P43411	methanococc
289	4	44.4	153	1	YPH4_CHRYI	P45373	chromatium		362	4	44.4	192	1	PM1_HUMAN	P17152	homo sapien
290	4	44.4	154	1	PUTR_RHOCA	O52710	rhodobacter		363	4	44.4	193	1	ASH2_HUMAN	O39929	homo sapien
291	4	44.4	155	1	BCCP_HAETN	P43874	haemophilus		364	4	44.4	193	1	HIS7_SUSO	O33773	sulfotobus
292	4	44.4	155	1	PRL_PHAVU	P25985	phaseolus v		365	4	44.4	194	1	PGHD_HORSE	O91921	equus cabal
293	4	44.4	155	1	PR2_PHAVU	P25986	phaseolus v		366	4	44.4	195	1	HPPK_SYNT3	P72736	synechocyst
294	4	44.4	156	1	CU55_ARADI	P80518	araneus dia		367	4	44.4	195	1	YI33_AGRU	P05681	agrobacteri
295	4	44.4	156	1	MAFK_HUMAN	O60675	homo sapien		368	4	44.4	196	1	AMEL_MOUSE	P45559	mus musculu
296	4	44.4	157	1	FLIN_PSEAE	O51466	pseudomonas		369	4	44.4	200	1	SSUE_PSEPU	O65762	pseudomonas
297	4	44.4	157	1	HS2C_CHURE	P12811	chlamydomon		370	4	44.4	201	1	COAT_LVX	P27334	illy virus
298	4	44.4	157	1	PRTC_CANPA	Q28278	canis famli		371	4	44.4	201	1	RB1B_RAT	P10536	rattus norv
299	4	44.4	157	1	YI17_YEAST	P40502	saccharomyc		372	4	44.4	201	1	YPT1_PHYTN	O01890	phytophthor
300	4	44.4	157	1	YO12_BPL2	P42547	bacterioph		373	4	44.4	202	1	AMEL_MONDO	Q28462	monodelphis
301	4	44.4	158	1	ILVH_BUCAP	O85294	buchnera ap		374	4	44.4	202	1	COAT_ELV	P14530	timelesuru
302	4	44.4	158	1	Y524_METJA	P26987	glycine max		375	4	44.4	202	1	HR2A_TRIFL	P14530	timelesuru
303	4	44.4	158	1	CUS7_ARADI	O57944	methanococc		376	4	44.4	202	1	RAB1_DISOM	P22125	discoyge o
304	4	44.4	159	1	YBAK_ECOLI	P80519	araneus dia		377	4	44.4	202	1	SDC4_RAT	P34901	rattus norv
305	4	44.4	159	1	YBAK_ECOLI	P31175	escherichia		378	4	44.4	204	1	T2M2_METTF	P29566	methanobact
306	4	44.4	159	1	YBAK_SALTY	P31174	salmonella		379	4	44.4	204	1	BFY3_MOUSE	O64152	mus musculu
307	4	44.4	160	1	PETD_ODOSI	P49489	odontella s		380	4	44.4	205	1	HYFA_ECOLI	P23481	escherichia
308	4	44.4	160	1	PHAA_CVAPA	P00316	cyranophora		381	4	44.4	205	1	RAB1_LYMET	O05974	lymaea sta
309	4	44.4	160	1	RL10_METTH	O27191	methanobact		382	4	44.4	205	1	RB1A_HUMAN	P11476	homo sapien
310	4	44.4	160	1	RL21_CAEEL	P34334	caenorhabdi		383	4	44.4	205	1	RB1A_RAT	P03329	rattus norv
311	4	44.4	161	1	ID4_MOUSE	P41139	mus musculu		384	4	44.4	206	1	COBH_METTH	O28462	methanobact
312	4	44.4	161	1	VG41_HSVSA	001027	herpesvirus		385	4	44.4	206	1	Y086_CAEEL	P34632	caenorhabdi
313	4	44.4	162	1	MAFG_HUMAN	O15525	homo sapien		386	4	44.4	207	1	PLCR_PSEAE	P40695	pseudomonas
314	4	44.4	162	1	MAFG_MOUSE	O54792	mus musculu		387	4	44.4	208	1	CSF3_MOUSE	P09920	mus musculu
315	4	44.4	163	1	HMCS_DICDI	P54872	dicyosteli		388	4	44.4	208	1	IL16_HORSE	O95181	equus cabal
316	4	44.4	163	1	ILVH_ECOLI	P00894	escherichia		389	4	44.4	208	1	NOLV_RHISN	P55717	rhizobium s
317	4	44.4	163	1	ILVH_HAETN	P45260	haemophilus		390	4	44.4	208	1	SPC3_STRPU	O48437	strongyloce
318	4	44.4	163	1	ILVH_SALTY	P21622	salmonella		391	4	44.4	209	1	YC02_KLEPN	P16548	klebsiella
319	4	44.4	164	1	FLAV_HELPJ	O94K53	helicobacte		392	4	44.4	209	1	YLZ3_CAEEL	P34416	caenorhabdi
320	4	44.4	164	1	FLAV_HELPJ	O25776	helicobacte		393	4	44.4	210	1	COBH_METJA	O58340	methanococc
321	4	44.4	164	1	IM17_SCHRO	P87130	schizosacch		394	4	44.4	210	1	R57_HALLA	P15763	halobacteri
322	4	44.4	164	1	NOLB_RHTR	P33208	rhizobium f		395	4	44.4	210	1	YI21_YEAST	P27654	saccharomyc
323	4	44.4	164	1	NOLB_RHISN	P55713	rhizobium s		396	4	44.4	210	1	TX10_HUMAN	O04133	homo sapien
324	4	44.4	164	1	TX2DA_XENLA	O91858	xenopus lae		397	4	44.4	211	1	DEDD_ECOLI	P09549	escherichia
325	4	44.4	164	1	TSTP_HSVSA	P18347	herpesvirus		398	4	44.4	211	1	TX10_HUMAN	O75333	homo sapien

399	4	44.4	212	1	YTFB_ECOLI	P39310	eschlerichia	472	4	44.4	244	1	YR01_CAEEL	Q10014	caenorhabd
400	4	44.4	213	1	CHRR_RHOSH	P40655	rhodobacter	473	4	44.4	245	1	KAD_CHLTR	084130	chlamydia t
401	4	44.4	213	1	HISS_LEGPN	O916X3	legionella	474	4	44.4	246	1	SIX6_HUMAN	095475	homo sapien
402	4	44.4	213	1	KITH_RHOSI	O92162	rhodothermu	475	4	44.4	246	1	Y984_CAMJE	P45491	campylobact
403	4	44.4	214	1	RSMB_RAT	P17136	rattus norv	476	4	44.4	247	1	CUT3_MYCTU	006318	mycobacteri
404	4	44.4	214	1	SC14_SCACO	P35755	schizophylli	477	4	44.4	247	1	PMX2_MOUSE	006348	mus muscucu
405	4	44.4	215	1	Y034_NPVAC	P21287	autographa	478	4	44.4	247	1	RS3_CAEEL	P48152	caenorhabd
406	4	44.4	217	1	YD67_SCIPO	O10319	schizosacch	479	4	44.4	247	1	SUMT_PSEFL	P37725	pseudomonas
407	4	44.4	219	1	H1B_XENLA	P06468	chlamydia t	480	4	44.4	248	1	TRYP_CLOAB	052693	clostridium
408	4	44.4	219	1	ISPD_CHLTR	O84468	chlamydia t	481	4	44.4	248	1	TRYP_FUSOX	P93206	fusarium ox
409	4	44.4	220	1	TRKA_MYCTU	O07194	mycobacteri	482	4	44.4	249	1	1431_LYCES	P93206	lycopersico
410	4	44.4	222	1	AMVB_SECCO	P30271	seccate cere	483	4	44.4	249	1	HMX1_CHICK	E28361	gallus gall
411	4	44.4	223	1	I12A_MARMO	O61728	marmota mon	484	4	44.4	249	1	REP2_ZYGBA	P13780	zygosacchar
412	4	44.4	224	1	DEOC_MERTH	O26909	methanobact	485	4	44.4	249	1	RU23_ARATH	P43323	arabidopsis
413	4	44.4	224	1	HXB6_HUMAN	P17509	homo sapien	486	4	44.4	249	1	YBFT_BACSU	O31458	bacillus su
414	4	44.4	224	1	HXB6_MOUSE	P09023	mus muscucu	487	4	44.4	250	1	FCG3_BOVIN	P79107	bos taurus
415	4	44.4	224	1	VV_MUMPE	P30927	mumps virus	488	4	44.4	250	1	PSDB_CAEEL	O23449	caenorhabd
416	4	44.4	224	1	VV_MUMPK	P30928	mumps virus	489	4	44.4	250	1	TAD2_YENST	P47058	saccharomyc
417	4	44.4	224	1	VV_MUMPS	P33483	mumps virus	490	4	44.4	250	1	YNCK_BACSU	O31603	bacillus su
418	4	44.4	224	1	YPJG_BACSU	P42981	bacillus su	491	4	44.4	251	1	CB24_ARATH	P27521	arabidopsis
419	4	44.4	225	1	XYNA_THELA	O43097	thermomyces	492	4	44.4	251	1	COBM_MYCTU	O10672	mycobacteri
420	4	44.4	226	1	ATP6_HELPJ	O92115	helicobacte	493	4	44.4	251	1	TPIS_LTIME	P48499	leishmania
421	4	44.4	226	1	ATP6_HELPY	P56085	helicobacte	494	4	44.4	251	1	TPIS_TRYCR	P52270	trypanosoma
422	4	44.4	226	1	CXB2_MOUSE	O00977	mus muscucu	495	4	44.4	251	1	TRPA_PYRKO	O9Y949	pyrococcus
423	4	44.4	226	1	CXB2_RAT	P21994	rattus norv	496	4	44.4	252	1	GSC_HUMAN	P56915	homo sapien
424	4	44.4	226	1	CXB2_SHEEP	P46691	ovis aries	497	4	44.4	253	1	EBSD_ENTFA	P36923	enterococcu
425	4	44.4	226	1	HMMB_BRARE	O01704	brachydanio	498	4	44.4	253	1	HBM4_MERTH	O26268	methanobact
426	4	44.4	226	1	LPOT_MYCTU	P96384	mycobacteri	499	4	44.4	253	1	KAD_CHLMU	O9KPD0	chlamydia m
427	4	44.4	227	1	SVT_PSEXY	P52833	pseudomonas	500	4	44.4	253	1	KDSB_HAEIN	P44900	haemophilus
428	4	44.4	228	1	RPE_RHOCA	P51012	rhodobacter	501	4	44.4	253	1	PMX2_HUMAN	O99811	homo sapien
429	4	44.4	229	1	IRLR_BURPS	O31395	burkholderi	502	4	44.4	254	1	GLTE_ECOLI	P28721	eschlerichia
430	4	44.4	229	1	Y4MB_RH1SN	P55561	rhizobium s	503	4	44.4	254	1	KDSB_CHLPN	O9Z8U9	chlamydia p
431	4	44.4	231	1	YD74_MYCPN	P75407	rhizobium s	504	4	44.4	254	1	YN08_YEAST	P53304	saccharomyc
432	4	44.4	231	1	FLGH_PSEPU	O52081	pseudomonas	505	4	44.4	255	1	GRPS_MYXXA	P95333	myxococcus
433	4	44.4	231	1	RSMB_MOUSE	P27048	mus muscucu	506	4	44.4	256	1	G1O2_RHOCA	P96881	rhodobacter
434	4	44.4	232	1	CASE_PIG	P39037	sus scrofa	507	4	44.4	256	1	GSC_MOUSE	O02591	mus muscucu
435	4	44.4	232	1	FLGH_ECOLI	P75940	eschlerichia	508	4	44.4	257	1	CBIF_SALTY	Q05630	salmonella
436	4	44.4	234	1	YG1B_ECOLI	P24195	eschlerichia	509	4	44.4	257	1	CG1C_ORYSA	P93411	oryza sativ
437	4	44.4	235	1	AURB_CHLAU	P27197	chloroflexu	510	4	44.4	257	1	COAT_TYLCM	P27257	tomato yell
438	4	44.4	235	1	MSRA_HUMAN	O9U168	homo sapien	511	4	44.4	257	1	COAT_TYLCU	P38608	tomato yell
439	4	44.4	235	1	VP26_VZVD	P09279	varicella-2	512	4	44.4	257	1	YZG1_CAEEL	P55226	caenorhabd
440	4	44.4	235	1	YHNO_HAEIN	P44908	haemophilus	513	4	44.4	258	1	HB21_SPAEH	P15464	spatial leuc
441	4	44.4	236	1	GAMT_HUMAN	O14353	homo sapien	514	4	44.4	258	1	HB2D_PIG	P15983	sus scrofa
442	4	44.4	236	1	GAMT_MOUSE	O35969	mus muscucu	515	4	44.4	259	1	MOTR_TREPA	O07886	treponema p
443	4	44.4	236	1	YMI3_CAEEL	O21018	caenorhabd	516	4	44.4	259	1	MSP8_EIMAC	P09128	elasmia ace
444	4	44.4	237	1	COX2_TIRIU	O01556	trichophyto	517	4	44.4	259	1	RNS2_ARATH	P42814	arabidopsis
445	4	44.4	237	1	DRNF_AERHY	O44064	aeromonas h	518	4	44.4	260	1	1433_LYCES	P93309	lycopersico
446	4	44.4	237	1	GUN_ASPAC	P22669	aspergillus	519	4	44.4	260	1	143C_TOBAC	P93343	nicotiana t
447	4	44.4	237	1	MTRA_METTH	O27227	methanobact	520	4	44.4	260	1	COAT_TYLCV	P27256	tomato yell
448	4	44.4	237	1	MTRA_METTM	P80184	methanobact	521	4	44.4	260	1	SHP_MOUSE	Q62227	mus muscucu
449	4	44.4	237	1	PSA5_SOYBN	O9M418	glycine max	522	4	44.4	260	1	VG38_BPK3	P07876	bacteriopho
450	4	44.4	237	1	YGHS_ECOLI	O46843	eschlerichia	523	4	44.4	261	1	CXHB_HUMAN	O95452	homo sapien
451	4	44.4	238	1	ET3_HUMAN	P14138	homo sapien	524	4	44.4	261	1	CXHB_MOUSE	P05689	mus muscucu
452	4	44.4	238	1	YT23_AGRVI	P70793	agrobacteri	525	4	44.4	261	1	HB22_HUMAN	P01919	homo sapien
453	4	44.4	239	1	RIM9_YEAST	O04734	saccharomyc	526	4	44.4	261	1	HB23_HUMAN	P03537	homo sapien
454	4	44.4	239	1	TRBH_ECOLI	P19381	eschlerichia	527	4	44.4	261	1	HB23_HUMAN	P03927	homo sapien
455	4	44.4	239	1	WT1_SMIMA	P48953	smithopsis	528	4	44.4	261	1	HB2C_PIG	P15982	sus scrofa
456	4	44.4	240	1	8512_TRYCR	P18270	trypanosoma	529	4	44.4	261	1	RNG6_MOUSE	P28078	mus muscucu
457	4	44.4	240	1	COAT_NMY	P15100	narcissus m	530	4	44.4	261	1	TFEP_MORBO	P20666	morexella b
458	4	44.4	240	1	FX24_RHITY	P24290	rhizobium l	531	4	44.4	261	1	Y497_MYCLE	P54880	mycobacteri
459	4	44.4	240	1	RSMB_CHICK	O9PV94	gallus gall	532	4	44.4	261	1	YF25_MYCTU	O50582	mycobacteri
460	4	44.4	240	1	RSMB_ERIEU	O9LUB7	erlinaceus e	533	4	44.4	261	1	YFJB_HAEIN	O44497	haemophilus
461	4	44.4	240	1	RSMB_HUMAN	P14678	homo sapien	534	4	44.4	261	1	YOB4_CAEEL	P09257	caenorhabd
462	4	44.4	240	1	RSMB_MONDO	O9LUB6	monodelphis	535	4	44.4	262	1	APAI_BRARE	O42363	brachydanio
463	4	44.4	240	1	RSMB_HUMAN	P14648	homo sapien	536	4	44.4	263	1	T2R8_THETH	P29278	thermus aqu
464	4	44.4	241	1	RASH_MSYHA	P01115	harvey muri	537	4	44.4	263	1	TRPA_ANTSP	P31204	antithamno
465	4	44.4	242	1	CCWC_RHOCA	P29961	rhodobacter	538	4	44.4	264	1	APAI_ANAPL	O42296	anas platyr
466	4	44.4	243	1	GSCA_XENLA	P29454	xenopus lae	539	4	44.4	264	1	APAI_CHICK	P08250	gallus gall
467	4	44.4	243	1	GSCB_XENLA	P53546	xenopus lae	540	4	44.4	264	1	APAI_COTJA	P32918	coturnix co
468	4	44.4	243	1	TRUA_MYCPN	O50291	mycoplasma	541	4	44.4	264	1	CBX1_XENLA	P08983	xenopus lae
469	4	44.4	243	1	VLPE_MYCHR	O49537	mycoplasma	542	4	44.4	264	1	DMPH_PSESP	P49156	pseudomonas
470	4	44.4	244	1	CAH_ERMCA	O52558	erythria car	543	4	44.4	264	1	GP3D_CHLMU	O46373	chlamydia m
471	4	44.4	244	1	H1O_CHLTH	O07134	chiltonomus	544	4	44.4	264	1	RSF4_STRPU	P46771	strongyloce

545	4	44.4	264	1	T2FB_XENLA	Q03123 xenopus lae	618	4	44.4	288	1	Y940_MYCTU	P71569 mycobacteri
546	4	44.4	264	1	TRPA_SYNN3	P77960 synechocyst	619	4	44.4	289	1	FA0E_SYNN3	P46017 anabena sp
547	4	44.4	264	1	VALL_MDV	P06847 wheat dwarf	620	4	44.4	290	1	AR0E_SYNN3	P74591 synechocyst
548	4	44.4	264	1	YMS3_CAEEL	Q10939 caenorhabdi	621	4	44.4	290	1	ITVE_RICPR	O05970 rickettsia
549	4	44.4	265	1	LECNI_PEA	P16270 pisum sativ	622	4	44.4	290	1	TRX2_MOUSE	O08550 mus musculu
550	4	44.4	265	1	PSBP_NARPS	Q40407 narctissus p	623	4	44.4	291	1	CC21_MEDSA	P24923 medicago sa
551	4	44.4	265	1	TRPA_MERTM	P26920 methanobact	624	4	44.4	291	1	CILB_HAEIN	P44460 haemophilus
552	4	44.4	265	1	YMI2_PARRE	P15612 paramecium	625	4	44.4	291	1	SP13_YEAST	P23624 saccharomyc
553	4	44.4	266	1	ELH2_ACICA	P00632 acinetobact	626	4	44.4	291	1	Y2EB_HUMAN	P29084 homo sapien
554	4	44.4	266	1	ETFB_MYCLE	Q33095 mycobacteri	627	4	44.4	291	1	YXUO_BACSU	P5181 bacillus su
555	4	44.4	266	1	HMUV_YERPE	Q56993 yerstinla pe	628	4	44.4	292	1	CC22_ORYSA	P29619 oryza sativ
556	4	44.4	266	1	SURE_MERTJA	Q57979 methanococ	629	4	44.4	292	1	Y185_BUCAI	P51282 buchnera ap
557	4	44.4	267	1	PSBP_SPIOL	P12302 spinnacia ol	630	4	44.4	293	1	HMX1_MOUSE	P31329 mus musculu
558	4	44.4	267	1	TRPA_BACSU	P07601 bacillus su	631	4	44.4	293	1	SYK3_AOUAE	O66963 aquifex aeo
559	4	44.4	268	1	INHA_MYCAV	O07400 mycobacteri	632	4	44.4	293	1	Y311_BORBU	O51291 borrellia bu
560	4	44.4	268	1	KSGA_RICPR	O05952 r dimethyla	633	4	44.4	294	1	CC21_ORYSA	P29618 oryza sativ
561	4	44.4	268	1	YHC6_YEAST	P38740 saccharomyc	634	4	44.4	294	1	CC22_MEDSA	O05006 medicago sa
562	4	44.4	269	1	INHA_MYCSM	P42829 mycobacteri	635	4	44.4	294	1	CC2A_ANTMA	Q38772 antirrhinum
563	4	44.4	269	1	INHA_MYCTU	P46533 mycobacteri	636	4	44.4	294	1	CC2A_ARATH	P24100 arabidopsis
564	4	44.4	269	1	NUDC_VIRCH	Q9VW27 vibrio chol	637	4	44.4	294	1	CC2_MATZE	P23111 zea mays (m
565	4	44.4	269	1	TRPA_BACST	P19867 bacillus st	638	4	44.4	294	1	CC2_VIGUN	P52389 vigna ungui
566	4	44.4	270	1	HCE1_ORYLA	P31580 oryzias lat	639	4	44.4	295	1	CAC3_HAEKO	P16283 haemochus
567	4	44.4	270	1	TRPA_MERTM	O27697 methanobact	640	4	44.4	295	1	YK05_MYCTU	Q10851 mycobacteri
568	4	44.4	271	1	Y228_TREPA	O83256 treponema p	641	4	44.4	296	1	CC01_CAEEL	P08124 caenorhabdi
569	4	44.4	272	1	RSR1_YEAST	P13856 saccharomyc	642	4	44.4	296	1	MY88_MOUSE	P22366 mus musculu
570	4	44.4	272	1	VG67_HSVB	P28984 equine hepr	643	4	44.4	297	1	CC2_BOVIN	P06493 bos taurus
571	4	44.4	272	1	YIS2_SHISO	P16940 shigella so	644	4	44.4	297	1	CC2_HUMAN	P06493 homo sapien
572	4	44.4	273	1	BASI_MOUSE	P18572 mus musculu	645	4	44.4	297	1	CC2_MOUSE	P11440 mus musculu
573	4	44.4	274	1	BPAL_STRAU	P33912 streptomyc	646	4	44.4	297	1	CC2_RAT	P39951 rattus norv
574	4	44.4	274	1	PRCE_ARATH	O23717 arabidopsis	647	4	44.4	297	1	HMX1_BOVIN	P02786 bos taurus
575	4	44.4	274	1	SAC7_YEAST	P17121 saccharomyc	648	4	44.4	297	1	HMX1_HUMAN	P28360 homo sapien
576	4	44.4	274	1	VGLE_HSV2	P31289 herpes simp	649	4	44.4	297	1	VDZC_SCHRO	Q13719 schistosach
577	4	44.4	274	1	YBIC_SCHRO	Q9URV0 schistosach	650	4	44.4	298	1	Y13K_HUMAN	O04323 homo sapien
578	4	44.4	274	1	YDDG_ECOLI	P46136 escherichia	651	4	44.4	299	1	PLTA_MYCPO	Q1263 mycosphaere
579	4	44.4	275	1	RK2_CYAPA	P15764 cyanophora	652	4	44.4	299	1	YC59_MYCTU	Q11059 mycobacteri
580	4	44.4	275	1	VE39_NPVAC	P11042 autographa	653	4	44.4	299	1	YH9_YEAST	P47032 saccharomyc
581	4	44.4	276	1	VGIM_PUUMB	P41264 pumala vir	654	4	44.4	300	1	CCPM_BACHD	Q92927 bacillus ha
582	4	44.4	276	1	COAT_TNVA	P22959 tobacco nec	655	4	44.4	301	1	ERA_BACSU	P10087 escherichia
583	4	44.4	276	1	RSPA_CAEEL	P46769 caenorhabdi	656	4	44.4	301	1	NHAR_ECOLI	P12423 rhododact
584	4	44.4	277	1	TRPA_HALVO	P18284 halobacteri	657	4	44.4	301	1	PORI_RHOCA	P31243 bacillus su
585	4	44.4	277	1	YAV4_PYRHO	O58801 pyrococcus	658	4	44.4	301	1	YXAC_BACSU	P42102 bacillus su
586	4	44.4	278	1	VGIL_HCMV1	O68667 human cytom	659	4	44.4	302	1	CASB_MACRO	P28550 macropus eu
587	4	44.4	278	1	VGIL_HCMW2	O68668 human cytom	660	4	44.4	302	1	CC21_XENLA	P33567 xenopus lae
588	4	44.4	278	1	VGIL_HCMW3	O68669 human cytom	661	4	44.4	302	1	CC22_XENLA	P24033 xenopus lae
589	4	44.4	278	1	VGIL_HCMW4	O68670 human cytom	662	4	44.4	302	1	CC2_CARAU	P51958 carassius a
590	4	44.4	278	1	VGIL_HCMW5	O68671 human cytom	663	4	44.4	302	1	CILB_ECOLI	P77770 escherichia
591	4	44.4	278	1	VGIL_HCMW6	O68672 human cytom	664	4	44.4	302	1	YE86_MYCTU	P71766 mycobacteri
592	4	44.4	278	1	VGIL_HCMW7	O68673 human cytom	665	4	44.4	303	1	CC2_CHICK	P13863 gallus gall
593	4	44.4	278	1	VGIL_HCMW8	O68674 human cytom	666	4	44.4	303	1	NIJM_CABUN	O78700 cabassous u
594	4	44.4	278	1	VGIL_HCMW9	P16832 human cytom	667	4	44.4	304	1	HEY1_CANFA	O95822 canis famli
595	4	44.4	279	1	VGIL_HCMWP	O68666 human cytom	668	4	44.4	304	1	HEY1_HUMAN	O95513 homo sapien
596	4	44.4	279	1	HCE2_ORYLA	P31581 oryzias lat	669	4	44.4	304	1	PRFB_BACAM	P41029 bacillus am
597	4	44.4	279	1	YA22_MERTM	O27101 methanobact	670	4	44.4	305	1	DDLB_ECOLI	O97860 escherichia
598	4	44.4	280	1	CC2B_ANTMA	Q38773 antirrhinum	671	4	44.4	305	1	HEM3_AERPE	O09910 aeropyrum p
599	4	44.4	280	1	Y747_RICPR	O9AC13 rickettsia	672	4	44.4	305	1	HM23_CAEEL	P34663 caenorhabdi
600	4	44.4	281	1	TRPA_MERTJA	O60180 methanococ	673	4	44.4	305	1	MAIM_SALTY	P26478 salmonella
601	4	44.4	281	1	YD3B_SCHRO	Q10275 schizosacch	674	4	44.4	305	1	Y40A_SYNN3	P74430 synechocyst
602	4	44.4	281	1	YESB_SCHRO	O14175 schizosacch	675	4	44.4	306	1	ARG1_BRUAB	O53174 bruceella ab
603	4	44.4	282	1	CHER_THEMA	Q9WY15 thermotoga	676	4	44.4	306	1	METF_STRLI	O54235 streptomyce
604	4	44.4	283	1	CXB1_HUMAN	P082034 homo sapien	677	4	44.4	307	1	RB33_CAEEL	O20363 caenorhabdi
605	4	44.4	283	1	CXB1_MOUSE	P28230 mus musculu	678	4	44.4	307	1	VP53_BPAPS	O39196 bacterioph
606	4	44.4	283	1	CXB1_RAT	P08033 rattus norv	679	4	44.4	307	1	Y955_MYCTU	O39196 mycobacteri
607	4	44.4	284	1	CRCY_CAPAE	Q29478 capra aegag	680	4	44.4	307	1	KDGD_BACSU	P42235 bacillus su
608	4	44.4	284	1	CK32_MITON	P51915 micropogoni	681	4	44.4	308	1	PRIB_MYCTU	O50563 mycobacteri
609	4	44.4	284	1	CXB1_BOVIN	O18968 bos taurus	682	4	44.4	308	1	TF2B_SULSH	O67619 aquifex aeo
610	4	44.4	284	1	RP32_PROMI	P16720 human cytom	683	4	44.4	310	1	APBA_AOUAE	P08222 mus musculu
611	4	44.4	284	1	UL06_HCMVA	O09378 caenorhabdi	684	4	44.4	311	1	PROC_NEUCR	O12641 neurospora
612	4	44.4	284	1	Y556_CAEEL	O09378 caenorhabdi	685	4	44.4	311	1	YBBU_BORBR	O06702 bordetella
613	4	44.4	288	1	FXE3_MOUSE	O99Y14 mus musculu	686	4	44.4	311	1	Y695_MYCLE	O49897 mycobacteri
614	4	44.4	288	1	HMX1_CHICK	P50223 gallus gall	687	4	44.4	311	1	FMT_THEMA	O9WY28 thermotoga
615	4	44.4	288	1	SMR2_PODAN	O08143 podospira a	688	4	44.4	313	1	LVTB_BURPS	O9FF10 burkholderi
616	4	44.4	288	1	WT1_AITMT	P50902 alligator m	689	4	44.4	313	1		
617	4	44.4	288	1			690	4	44.4	313	1		

691	4	44.4	313	1	PEPM_STRHY	P29247	streptomyc	764	4	44.4	332	1	LDHH_CHICK	P00337	gallus gall
692	4	44.4	313	1	VU47_HSVJ7	P52555	human herpe	765	4	44.4	332	1	LDH_DROME	095028	drosophila
693	4	44.4	314	1	ACCD_ANGLY	P28252	angliopteris	766	4	44.4	332	1	LYTB_MYCLE	09x781	mycobacteri
694	4	44.4	314	1	FMT_PSEAE	085732	pseudomonas	767	4	44.4	332	1	SIX3_HUMAN	095333	homo sapien
695	4	44.4	314	1	SIX3_CHICK	042406	gallus gall	768	4	44.4	332	1	SR4_PHYTO	P11113	physarum po
696	4	44.4	314	1	YM78_YEAST	004013	saccharomyc	769	4	44.4	332	1	THCR_RHOER	P43462	rhodococcus
697	4	44.4	315	1	GBF7_ARATH	P42774	arabidopsis	770	4	44.4	333	1	SIX3_MOUSE	062223	mus musculu
698	4	44.4	315	1	HYFC_ECOLI	P72858	escherichia	771	4	44.4	333	1	VINT_BPMD2	038381	mycobacteri
699	4	44.4	315	1	ISPE_SYNY3	P72663	synecocyst	772	4	44.4	333	1	VINT_BPMFR	P25426	mycobacteri
700	4	44.4	315	1	MCH_METKA	P49654	methanopyru	773	4	44.4	334	1	CXA2_XENLA	P16864	xenopus lae
701	4	44.4	315	1	RPOA_CLOPE	091bw9	clostridium	774	4	44.4	334	1	MRPL_CAEEL	093439	caenorhabdi
702	4	44.4	316	1	MCH_ARCEU	028344	archaeoglob	775	4	44.4	334	1	YI11_HALNT	P23464	halobacteri
703	4	44.4	316	1	VAM7_YEAST	P32912	saccharomyc	776	4	44.4	335	1	FIMH_SALTY	P37925	salmonella
704	4	44.4	317	1	APE_HUMAN	P02649	homo sapien	777	4	44.4	335	1	G3P_HALVA	054388	halocaula
705	4	44.4	317	1	FLGL_ECOLI	P29744	escherichia	778	4	44.4	335	1	LYB1_MYCTU	P38200	mycobacteri
706	4	44.4	317	1	LDH_BACCA	P10655	bacillus ca	779	4	44.4	335	1	SLAM_HUMAN	013291	homo sapien
707	4	44.4	317	1	RSP4_URECA	P38981	urechis cau	780	4	44.4	335	1	T2G1_HERAU	P25261	herpetosiph
708	4	44.4	318	1	ATPS_YEAST	P32453	saccharomyc	781	4	44.4	337	1	FABH_PSEAE	P20582	pseudomonas
709	4	44.4	318	1	CH11_HORVU	P11955	hordium vul	782	4	44.4	337	1	G3PC_MESCR	P17878	mesembryant
710	4	44.4	318	1	NUIM_TAMTE	078699	lamandua te	783	4	44.4	337	1	G3P_LACIA	P52987	lactococcus
711	4	44.4	318	1	ONCA_TAETA	P22080	taenia taen	784	4	44.4	337	1	KCIA_CHICK	P70065	gallus gall
712	4	44.4	319	1	FMT_TREPA	083737	treponema p	785	4	44.4	337	1	KCIA_HUMAN	P48729	homo sapien
713	4	44.4	319	1	HA11_RAT	P15978	rattus norv	786	4	44.4	337	1	YFGA_ECOLI	P27424	escherichia
714	4	44.4	319	1	MCH_METTM	P51616	methanobact	787	4	44.4	338	1	MBR3_YEAST	P32488	saccharomyc
715	4	44.4	319	1	MOCB_STYP7	Q56208	synecococc	788	4	44.4	338	1	TR19_HUMAN	015653	homo sapien
716	4	44.4	320	1	GSHB_SYNY3	Q01578	synecococ m	789	4	44.4	338	1	YBD1_YEAST	P19304	saccharomyc
717	4	44.4	320	1	HXA4_HUMAN	000056	homo sapien	790	4	44.4	339	1	KDGT_ERWCH	P15701	erynia chr
718	4	44.4	320	1	MCH_METTH	026867	methanobact	791	4	44.4	339	1	PCB3_HUMAN	P57721	homo sapien
719	4	44.4	320	1	THTR_STYP7	P27477	synecococc	792	4	44.4	339	1	PYRD_HAETN	P45477	haemophilus
720	4	44.4	321	1	LIPA_ECOLI	P25845	escherichia	793	4	44.4	339	1	RLJ29_SPLIC	031163	spiroplasma
721	4	44.4	321	1	MCH_METBA	P94919	methanosarc	794	4	44.4	339	1	YJGB_ECOLI	P27220	escherichia
722	4	44.4	321	1	SAPB_ECOLI	Q47663	escherichia	795	4	44.4	339	1	YNZ8_CAEEL	P49568	caenorhabdi
723	4	44.4	321	1	SAPB_SALTY	P36628	salmonella	796	4	44.4	340	1	DAF_PONPY	P49457	pomo pygma
724	4	44.4	321	1	Y541_PVRHO	058276	pyrococcus	797	4	44.4	340	1	YMD7_YEAST	003703	saccharomyc
725	4	44.4	322	1	ADT_SCHRO	Q09188	schizosacch	798	4	44.4	340	1	COA2_POVMK	P24596	mouse polyo
726	4	44.4	322	1	HA10_MOUSE	Q01898	mus musculu	799	4	44.4	341	1	MDHW_CAEEL	002640	caenorhabdi
727	4	44.4	322	1	HCR_ECOLI	P75824	escherichia	800	4	44.4	341	1	HEAD_BP21	P03020	bacterioph
728	4	44.4	322	1	FLK_RAT	P09760	rattus norv	801	4	44.4	342	1	GALR_ECOLI	P03024	escherichia
729	4	44.4	323	1	MCH_METEX	085014	methylobact	802	4	44.4	343	1	VMAT_SSPVB	P16628	subcutis sc
730	4	44.4	323	1	MCH_METJA	059030	methanococc	803	4	44.4	344	1	RECA_MAGMG	P16034	magnetospir
731	4	44.4	323	1	OTX1_BRARE	Q01994	brachydanio	804	4	44.4	344	1	AOX_TRYBB	026710	trypanosoma
732	4	44.4	323	1	THIG_RHET	Q34293	thizobium e	805	4	44.4	345	1	ODPA_ACHLA	P34505	acholeplasm
733	4	44.4	323	1	THIG_RHET	Q05668	mycobacteri	806	4	44.4	345	1	VE2_HP34	P36792	human papil
734	4	44.4	323	1	YM98_MYCTU	P14012	agrobacteri	807	4	44.4	346	1	HRP_PSESY	005395	pseudomonas
735	4	44.4	324	1	ARG1_AGRTO	P25916	mus musculu	808	4	44.4	346	1	KORA_ARCFU	029781	archaeoglob
736	4	44.4	324	1	BM11_MOUSE	P40880	hordium vul	809	4	44.4	346	1	NU2M_CONJA	P24971	coltrinx co
737	4	44.4	324	1	CAHC_HORVU	P35506	bos taurus	810	4	44.4	346	1	YI61_XANCA	P25478	xanthomonas
738	4	44.4	325	1	KCIA_BOVIN	P97633	rattus norv	811	4	44.4	347	1	GABC_HUMAN	006545	homo sapien
739	4	44.4	325	1	BM11_HUMAN	P35226	homo sapien	812	4	44.4	347	1	GABC_MOUSE	000421	mus musculu
740	4	44.4	326	1	TH14_YEAST	P32318	saccharomyc	813	4	44.4	347	1	VIT3_CHICK	091025	gallus gall
741	4	44.4	326	1	VP40_EBOZM	Q05158	ebola virus	814	4	44.4	347	1	EXOB_AZOBK	059083	azospirillum
742	4	44.4	326	1	VS09_ROTHT	P10501	human rotav	815	4	44.4	348	1	NU2M_ONCMY	P48175	oncorhynch
743	4	44.4	326	1	VS09_ROTHT	P11865	human rotav	816	4	44.4	349	1	NU2M_SALSA	035924	salmo salar
744	4	44.4	326	1	VS09_ROTHT	P09365	porcine rot	817	4	44.4	349	1	STY1_SCHPO	009882	schizosacch
745	4	44.4	327	1	CP27_PIG	P79402	sus scrofa	818	4	44.4	349	1	VAL1_PHOV	006923	pepper huas
746	4	44.4	327	1	ENT2_MOUSE	Q61672	m equilibra	819	4	44.4	351	1	CX41_XENLA	P51914	xenopus lae
747	4	44.4	327	1	GLSA_SYNY3	P73905	synecocyst	820	4	44.4	351	1	MSX2_YEAST	P04990	saccharomyc
748	4	44.4	327	1	HMDL_DROME	P20009	drosophila	821	4	44.4	351	1	MURR_BORPE	09x628	bordeletia
749	4	44.4	328	1	COPB_PRESM	P12375	pseudomonas	822	4	44.4	351	1	PSBD_GUTTH	078427	guillartidia
750	4	44.4	328	1	GNDS_HUMAN	Q12967	homo sapien	823	4	44.4	351	1	PSBD_CHIVU	P49478	odontella s
751	4	44.4	328	1	IBP2_HUMAN	086779	homo sapien	824	4	44.4	351	1	PSBD_ODOST	P24090	ratuys norv
752	4	44.4	329	1	COAA_STRGO	Q11158	mycobacteri	825	4	44.4	352	1	AZHS_RAT	P50863	baclillus su
753	4	44.4	329	1	Y493_MYCTU	095154	homo sapien	826	4	44.4	352	1	MRP_BACSU	P06007	chlamydomon
754	4	44.4	331	1	AR73_HUMAN	P56670	mesocricetu	827	4	44.4	352	1	PSBD_CHIRE	P56319	chlorella v
755	4	44.4	331	1	MAZ_MESNU	059599	colletotric	828	4	44.4	352	1	PSBD_CYAPA	P48079	cyanophora
756	4	44.4	331	1	PELB_COLGL	P79019	aspergillus	829	4	44.4	352	1	PSBD_EUGGR	P06006	pisum sativ
757	4	44.4	331	1	AXHA_ASPNG	P79021	aspergillus	830	4	44.4	352	1	PSBD_SPLOL	P11005	synecococc
758	4	44.4	332	1	AXHA_ASPNG	P79021	aspergillus	831	4	44.4	352	1	TPO_CANFA	P42705	canis famli
759	4	44.4	332	1	AXHA_ASPNG	P79021	aspergillus	832	4	44.4	352	1	CHLI_ODOST	P49469	odontella s
760	4	44.4	332	1	AXHA_ASPNG	P79021	aspergillus	833	4	44.4	352	1			
761	4	44.4	332	1	AXHA_ASPNG	P79021	aspergillus	834	4	44.4	352	1			
762	4	44.4	332	1	AXHA_ASPNG	P79021	aspergillus	835	4	44.4	352	1			
763	4	44.4	332	1	AXHA_ASPNG	P79021	aspergillus	836	4	44.4	352	1			

837	4	44.4	353	1	LEU3_BACFR	P54354 bacteroides	910	4	44.4	370	1	KCC1_HUMAN	Q14012 homo sapien
838	4	44.4	353	1	PSBD_ARATH	P56167 arabidopsis	911	4	44.4	370	1	KDPO_SYNY3	P73870 synechocyst
839	4	44.4	353	1	PSBD_HORVU	P11849 hordeum vul	912	4	44.4	370	1	ODPA_BACSU	P21881 bacillus su
840	4	44.4	353	1	PSBD_MAIZE	P48184 zea mays (m	913	4	44.4	370	1	VP40_BDV	Q01552 borna disea
841	4	44.4	353	1	PSBD_MARPO	P06404 marchantia	914	4	44.4	371	1	HA12_RAT	P16391 rattus norv
842	4	44.4	353	1	PSBD_ORYSA	P12095 oryza sativ	915	4	44.4	371	1	MURG_SYNY3	P74657 synechocyst
843	4	44.4	353	1	PSBD_PINTH	P41644 pinus thunb	916	4	44.4	371	1	VZRH_HUMAN	P30518 homo sapien
844	4	44.4	353	1	PSBD_SECEH	P10803 secale cere	917	4	44.4	371	1	YOJE_BACSU	P54542 bacillus su
845	4	44.4	353	1	PSBD_TOBAC	P06403 nicotiana t	918	4	44.4	372	1	HMBN_BOMMO	P27605 bombyx mori
846	4	44.4	354	1	CVE2_AGRVU	P54082 agrobacteri	919	4	44.4	372	1	NOIC_SYNY3	P26522 synechocyst
847	4	44.4	354	1	MTD2_HERAU	P25265 herpetosiph	920	4	44.4	373	1	MLE_TRICU	P46057 trichosporo
848	4	44.4	354	1	Y415_SYNY3	P22040 synechocyst	921	4	44.4	373	1	YN28_YEAST	P53829 saccharomyc
849	4	44.4	355	1	CXA5_RAT	P28234 rattus norv	922	4	44.4	374	1	ADH3_KLUJA	P49384 kluyveromyc
850	4	44.4	356	1	CXA5_CANFA	P33725 canis famli	923	4	44.4	374	1	FUT5_HUMAN	O11128 homo sapien
851	4	44.4	356	1	NTRB_RHOCA	P09431 rhodobacter	924	4	44.4	374	1	FUT5_PANTR	P56433 pan troglod
852	4	44.4	356	1	TPO_MOUSE	P40226 mus musculu	925	4	44.4	374	1	O71A_DROME	Q9VUK5 drosophila
853	4	44.4	357	1	CXA5_HUMAN	P36382 homo sapien	926	4	44.4	374	1	FLIG_BPT4	P00971 bacterioph
854	4	44.4	357	1	CXA5_MOUSE	O01231 mus musculu	927	4	44.4	375	1	FRS2_HALN1	Q48290 halobacteri
855	4	44.4	357	1	PYRD_MYCTU	O06236 mycobacteri	928	4	44.4	375	1	YKUS_CAEEL	O17778 caenorhabdi
856	4	44.4	358	1	AZHS_CAVPO	O70159 cavia porce	929	4	44.4	376	1	BCA2_YEAST	P41399 botdetella
857	4	44.4	358	1	LEU3_HAETN	P43860 haemophilus	930	4	44.4	376	1	DHAS_BOREP	O84198 thermus aqu
858	4	44.4	358	1	TTTC3_AGRVI	O34296 agrobacteri	931	4	44.4	376	1	HOSC_THERH	P38779 saccharomyc
859	4	44.4	358	1	TTTC4_AGRVI	P70792 agrobacteri	932	4	44.4	376	1	YHL2_YEAST	P16863 xenopus lae
860	4	44.4	358	1	TTTC5_AGRVI	O34295 agrobacteri	933	4	44.4	378	1	CXAL_XENLA	P29303 drosophila
861	4	44.4	358	1	VCOM_ADE40	P48753 human adeno	934	4	44.4	378	1	HAIR_DROYI	P36790 human papil
862	4	44.4	359	1	GUN_CELUD	P18336 cellulomona	935	4	44.4	378	1	VE2_HPV30	P45867 bacillus su
863	4	44.4	359	1	MANR_PSEPU	P14444 pseudomonas	936	4	44.4	379	1	ACDA_BACSU	P28004 saccharomyc
864	4	44.4	359	1	T2D1_HERAU	P24599 herpetosiph	937	4	44.4	379	1	YAD2_YEAST	P28004 saccharomyc
865	4	44.4	359	1	YRA7_CAEEL	P34258 caenorhabdi	938	4	44.4	380	1	COSC_YEAST	P53053 saccharomyc
866	4	44.4	360	1	MDH_PYRHO	O59028 pyrococcus	939	4	44.4	380	1	CXAL_BRARE	O54474 brachydanio
867	4	44.4	360	1	PSB1_SYNP7	P04996 synechococc	940	4	44.4	381	1	DAF_HUMAN	P08174 homo sapien
868	4	44.4	360	1	PSB2_SYNP7	P04997 synechococc	941	4	44.4	381	1	Y926_HELPY	P55985 helicobacte
869	4	44.4	360	1	UPP_TREPA	O83462 leptonema p	942	4	44.4	381	1	YBD2_YEAST	P38199 saccharomyc
870	4	44.4	360	1	VC47_BPMO	O911v2 bacterioph	943	4	44.4	382	1	COAL_POVNA	P03020 mouse polyo
871	4	44.4	361	1	COBT_MYCTU	O10396 mycobacteri	944	4	44.4	382	1	GABB_MOUSE	O00420 mus musculu
872	4	44.4	361	1	VCOM_ADEB2	O96625 bovine aden	945	4	44.4	382	1	MYB_AV1MB	P01104 avian myelo
873	4	44.4	362	1	HA11_MOUSE	P01899 mus musculu	946	4	44.4	382	1	VMAT_P14HA	P27019 human para
874	4	44.4	362	1	HA11_MOUSE	P01897 mus musculu	947	4	44.4	382	1	VMAT_P14HB	P27020 human para
875	4	44.4	363	1	MURG_BORBU	O51708 botrellia bu	948	4	44.4	383	1	COAL_POVNA3	P03091 mouse polyo
876	4	44.4	364	1	MURG_STRCO	O92bas streptomyc	949	4	44.4	383	1	COAL_POVNC	P12907 mouse polyo
877	4	44.4	364	1	RECFF_AZOV1	P49997 azotobacteri	950	4	44.4	383	1	COAL_POVMP	P49302 mouse polyo
878	4	44.4	364	1	TTTC1_AGRVI	O44497 agrobacteri	951	4	44.4	383	1	VE2_HPV03	P36778 human papil
879	4	44.4	364	1	TTTC2_AGRVI	P70787 agrobacteri	952	4	44.4	384	1	VE2_HPV53	P36797 human papil
880	4	44.4	365	1	COAL_POVRO	P24848 bovine poly	953	4	44.4	385	1	CD34_HUMAN	P28906 homo sapien
881	4	44.4	365	1	COMT_MEDSA	P28002 medicago sa	954	4	44.4	385	1	YD16_SCHPO	O92344 schizosacch
882	4	44.4	365	1	COMT_POPTM	O00763 populus tre	955	4	44.4	385	1	YH1U_ECOCI	P71636 escherichia
883	4	44.4	365	1	COMT_PPRDU	O43609 prunus dulc	956	4	44.4	386	1	PSD4_ARATH	P55034 arabidopsis
884	4	44.4	365	1	ECT1_SCHPO	O9utic6 schizosacch	957	4	44.4	387	1	NMT_DROME	O61613 drosophila
885	4	44.4	365	1	GLDA_PSEPU	P50173 pseudomonas	958	4	44.4	387	1	U133_RCMVM	O12000 rat cytoleg
886	4	44.4	365	1	NEMA_ECOCI	P77256 escherichia	959	4	44.4	387	1	VE2_HPV41	P27552 human papil
887	4	44.4	366	1	RECFF_PSEPU	P13456 pseudomonas	960	4	44.4	390	1	ACKA_MYCPN	P75245 mycoplasma
888	4	44.4	366	1	CYCR_RHOGE	P51758 rhodocyclus	961	4	44.4	390	1	IDA_A_ECOCI	P33377 escherichia
889	4	44.4	366	1	OM3A_RHILV	O05811 rhizobium 1	962	4	44.4	390	1	P53_MOUSE	P03340 mus musculu
890	4	44.4	367	1	AZHS_HUMAN	P02765 homo sapien	963	4	44.4	390	1	RBP_MUMP1	P19717 mumps virus
891	4	44.4	367	1	CRPT_CRIGR	P49584 cricetulus	964	4	44.4	391	1	DPB_SYNY3	P72856 synechocyst
892	4	44.4	367	1	CRPT_HUMAN	P49585 homo sapien	965	4	44.4	391	1	HERP_MOUSE	O911k5 mus musculu
893	4	44.4	367	1	CRPT_MOUSE	P49586 mus musculu	966	4	44.4	391	1	KC21_CHICK	P21868 gallus galli
894	4	44.4	367	1	CRPT_RAT	P19836 rattus norv	967	4	44.4	391	1	KC21_HUMAN	P16138 homo sapien
895	4	44.4	367	1	GOX1_ARATH	O91irs0 arabidopsis	968	4	44.4	391	1	KC21_MOUSE	O60737 mus musculu
896	4	44.4	367	1	GOX2_ARATH	P16926 escherichia	969	4	44.4	391	1	KC21_RABIT	P33674 oryctolagus
897	4	44.4	367	1	MREC_ECOCI	P16926 escherichia	970	4	44.4	391	1	KC21_RAT	P19139 rattus norv
898	4	44.4	367	1	TTSD_MOUSE	P23949 mus musculu	971	4	44.4	391	1	P53_RAT	P10361 rattus norv
899	4	44.4	368	1	CXA5_CHICK	P18860 gallus galli	972	4	44.4	391	1	RBP_MUMPE	P16072 mumps virus
900	4	44.4	368	1	HA14_MOUSE	P14427 mus musculu	973	4	44.4	391	1	RBP_MUMPM	P16595 mumps virus
901	4	44.4	368	1	HA1D_MOUSE	P01902 mus musculu	974	4	44.4	391	1	VIE2_MCNVS	P24909 murine cyto
902	4	44.4	368	1	HA1M_MOUSE	P03991 mus musculu	975	4	44.4	392	1	SYD1_RAT	O55145 rattus norv
903	4	44.4	368	1	LIVE_MYCTU	O10399 mycobacteri	976	4	44.4	393	1	C143_MYCTU	O59336 mycobacteri
904	4	44.4	368	1	ODPA_BACST	P21873 bacillus st	977	4	44.4	393	1	DSOR_DROME	O24324 drosophila
905	4	44.4	368	1	VE2_HPV6A	O84294 human papil	978	4	44.4	393	1	LA2M_MYCSM	P21795 mycobacteri
906	4	44.4	368	1	YC08_SULSO	P95997 sulfolobus	979	4	44.4	393	1	P53_HUMAN	P04637 homo sapien
907	4	44.4	369	1	Y11_MOUSE	O62481 mus musculu	980	4	44.4	393	1	P53_TUPGB	O9ctai1 tupala glis
908	4	44.4	369	1	HA1B_MOUSE	P01901 mus musculu	981	4	44.4	393	1	Y1P3_YEAST	P4454 saccharomyc
909	4	44.4	369	1	HA1K_MOUSE	P04223 mus musculu	982	4	44.4	394	1	A2AB_RABIT	O77830 oryctolagus

```

983 4 44.4 394 1 CC91_YEAST P41733 saccharomyc
984 4 44.4 394 1 NUCC_SYNY3 P27724 synechocyst
985 4 44.4 394 1 PGK_BACSU P40924 bacillus su
986 4 44.4 394 1 Y332_SCHPO 009709 schizosacch
987 4 44.4 394 1 YEDE_ECOLI P76518 escherichia
988 4 44.4 395 1 HMPA_ERWCH Q47266 erwilia chr
989 4 44.4 395 1 LEUK_MOUSE P15702 mus musculu
990 4 44.4 395 1 LHX3_XENTIA P36200 xenopus lae
991 4 44.4 396 1 AGS_AGRRH P27875 agrobacteri
992 4 44.4 396 1 DCOP_SORMA P78748 sordaria ma
993 4 44.4 396 1 EPTU_STRAU P42479 stigmatella
994 4 44.4 396 1 GLPC_ECOLI P13034 escherichia
995 4 44.4 396 1 YJTB_BACSU 034374 bacillus su
996 4 44.4 397 1 YL65_MYCTU 006212 mycobacteri
997 4 44.4 397 1 DCOP_NEUCR P05035 neurospora
998 4 44.4 397 1 STAD_GOSHI 042770 gossypium h
999 4 44.4 398 1 TAP_DROME 016867 drosophila
1000 4 44.4 399 1 CXAB_CHICK P36381 gallus gall

```

## ALIGNMENTS

```

RESULT 1
P60_LISMO STANDARD; PRT; 484 AA.
ID P60_LISMO
AC P21171;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROTEIN P60 PRECURSOR (INVASION-ASSOCIATED PROTEIN).
GN IAP.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A. AND SEQUENCE OF 28-49.
RC STRAIN-SV1/2A BGD(KAUFMANN);
RX MEDLINE=90256283; Pubmed=2111287;
RA Koehler S., Leimeister-Waechter M., Chakraborty T., Lottspeich F.,
RA Goebel W.;
RT "The gene coding for protein p60 of Listeria monocytogenes and its
RT use as a specific probe for Listeria monocytogenes.";
RL Infect. Immun. 58:1943-1950(1990).
RN [2]
RP MEDLINE=93094153; Pubmed=1459966;
RX Buber A., Kuhn W., Goebel W., Koehler S.;
RT "Structural and functional properties of the p60 proteins from
RT different Listeria species.";
RL J. Bacteriol. 174:8166-8171(1992).
CC -1- FUNCTION: THIS MAJOR EXTRACELLULAR PROTEIN MAY BE INVOLVED IN THE
CC INVASION OF NONPROFESSIONAL PHAGOCYTIC CELLS BY LISTERIA.
CC -1- SIMILARITY: BELONGS TO THE E. COLI NLPC / LISTERIA P60 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X52268; CAA36509.1; -
CC PIR: A41487; A41487.
CC InterPro: IPR000064; -
CC DR InterPro: IPR002482; -
CC Pfam: PF00877; NLPC_P60; 1.
CC Pfam: PF01476; PG_blding_2; 2.
CC Signal: Repeat.
CC FT SIGNAL 1 27
CC CHAIN 28 484 PROTEIN P60.

```

```

FT DOMAIN 311 355 19 X 2 AA TANDEM REPEATS OF T-N.
SQ SEQUENCE 484 AA; 50587 MW; 3C0F90591E14E0F CRC64;

```

```

Query Match 100.0%; Score 9; DB 1; Length 484;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VSTPVAPPTQ 9
Db 145 VSTPVAPPTQ 153

```

## RESULT 2

```

PRSL_SOLTV STANDARD; PRT; 155 AA.
ID PRSL_SOLTV
AC P17641;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)
DE PATHOGENESIS-RELATED PROTEIN STH-21.
GN STH-21.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; easterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91346678; Pubmed=2102864;
RA Matton D.P., Brisson N.;
RT "Nucleotide sequence of a pathogenesis-related gene of potato.";
RL Plant Mol. Biol. 14:863-865(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92190624; Pubmed=2520162;
RA Matton D.P., Brisson N.;
RT "Cloning, expression, and sequence conservation of
RT pathogenesis-related gene transcripts of potato.";
RL Mol. Plant Microbe Interact. 2:325-331(1989).
CC -1- INDUCTION: BY WOUNDING AND ELICITOR TREATMENTS.
CC -1- SIMILARITY: BELONGS TO THE BETV1 FAMILY OF PATHOGENESIS-RELATED
CC PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M25156; AAA03020.1; -
CC DR EMBL: M28042; AAA02829.1; -
CC PIR: S11869; S11869.
CC DR HSSP: P15494; 1BTV.
CC DR InterPro: IPR000916; -
CC DR Pfam: PF00407; Bet.v.I.1.
CC DR PRINTS: PR00634; BETALLERGEN.
CC DR PROSITE: PS00451; PATHOGENESIS-BETV1; 1.
CC KW Pathogenesis-related protein; Multigene family.
SQ SEQUENCE 155 AA; 17200 MW; 02A78EB6D26A7CD9 CRC64;

```

```

Query Match 66.7%; Score 6; DB 1; Length 155;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 TPVAPT 8
Db 12 TPVAPT 17

```

```

RESULT 3
VCG_BPPHK STANDARD: PRT: 187 AA.
AC 038042:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE MAJOR SPIKE PROTEIN (G PROTEIN) (SPG).
GN G.
OS Bacteriophage phi-K.
OC Viruses: ssDNA viruses; Microviridae; Microvirus.
ON NCBI_TaxID=10848;
RX [1]
RP SEQUENCE FROM N.A.
RA Kodaira K.-I., Tadokoro H., Suzuki K., Oki M., Manto S., Taketo A.;
RU Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR SPIKE COMPONENT. INVOLVED IN THE ATTACHMENT TO
CC THE BACTERIAL HOST.
CC -1- SUBUNIT: THE VIRION IS COMPOSED OF 60 COPIES EACH OF THE F, G, AND
CC J PROTEINS, AND 12 COPIES OF THE H PROTEIN. THERE ARE 12 SPIKES
CC WHICH ARE EACH COMPOSED OF 5 G AND ONE H PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
-----
CC EMBL: X60323; CAA42892.1; -.
CC DR HSSP: P03643; IALU.
CC KW Coat protein
SQ SEQUENCE 187 AA: 19569 MW: 35C5CEC891241701 CRC64;

```

```

Query Match 66.7%; Score 6; DB 1; Length 187;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 TPVAPT 8
Db 25 TPVAPT 30

```

```

RESULT 4
SOD3_PLEBO STANDARD: PRT: 239 AA.
AC P50060:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SUPEROXIDE DISMUTASE [MN] 3 PRECURSOR (EC 1.15.1.1) (FRAGMENT).
GN SOD3.
OS Plectonema boryanum.
OC Bacteria: Cyanobacteria; Oscillatoriales; Plectonema.
ON NCBI_TaxID=1184;
RX [1]
RP SEQUENCE FROM N.A.
RA STRAIN-UTEX 485;
RX MEDLINE-95164530; PubMed-7860607;
RA Campbell W.S., Landenbach D.E.;
RT "Characterization of four superoxide dismutase genes from a
RT filamentous cyanobacterium."
J. Bacteriol. 177:964-972(1995).
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) -> O(2) + H(2)O(2).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
-----
CC EMBL: U17611; AAA69953.1; -.
CC DR HSSP: P09214; LMNG.
CC DR InterPro: IPR001189; -.
CC DR Pfam: PF00081; sodfe; 1.
CC DR PROSITE: PS00088; SOD_MN; 1.
CC KW Oxidoreductase; Manganese; Multigene family; Signal.
FT NON_TER 1 1
FT SIGNAL <1 30 POTENTIAL.
FT CHAIN 31 239 SUPEROXIDE DISMUTASE [MN] 3.
FT METAL 57 57 MANGANESE(BY SIMILARITY).
FT METAL 112 112 MANGANESE(BY SIMILARITY).
FT METAL 195 195 MANGANESE(BY SIMILARITY).
FT METAL 199 199 MANGANESE(BY SIMILARITY).
SQ SEQUENCE 239 AA: 26854 MW: 242CBD3C9C61D100 CRC64;

```

```

Query Match 66.7%; Score 6; DB 1; Length 239;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VSTPVA 6
Db 17 VSTPVA 22

```

```

RESULT 5
GATA_HELPJ STANDARD: PRT: 453 AA.
ID GATA_HELPJ
AC Q9ZL13;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE SUBUNIT A (EC 6.3.5.-) (GLU-ADT
DE SUBUNIT A).
GN GATA OR JHP0769.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
ON NCBI_TaxID=85963;
RX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99120557; PubMed-9923682;
RA Alm R.A., Ling L.S.L., Molt D.T., King B.L., Brown E.D., Dolg P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
Nature 397:176-180(1999).
CC -1- FUNCTION: FURNISHES A MEANS FOR FORMATION OF CORRECTLY CHARGED
CC GLN-TRNA(GLN) THROUGH THE TRANSAMINATION OF MISACYLATED GLU-
CC TRNA(GLN) IN ORGANISMS WHICH LACK GLUTAMINYL-TRNA SYNTHETASE. THE
CC REACTION TAKES PLACE IN THE PRESENCE OF GLUTAMINE AND ATP THROUGH
CC AN ACTIVATED GAMMA-PHOSPHO-GLU-TRNA(GLN) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMYL-TRNA(GLN) + L-GLUTAMINE -> ADP
CC + PHOSPHATE + L-GLUTAMINYL-TRNA(GLN) + L-GLUTAMATE.
CC -1- SUBUNIT: HETEROTRIMER OF A, B AND C SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/

```



CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch).  
CC -----  
CC EMBL: AE001507; AAD06348.1; -  
CC InterPro: IPR000120; -  
CC Pfam: PF01425; AMIDase; 1.  
CC PROSITE: PS00571; AMIDASES; 1.  
CC KEGG: Protein biosynthesis; Ligase.  
CC SEQUENCE 453 AA; 49749 MW; FD97731532E77397 CRC64;  
  
Query Match  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 TPVAPT 8  
|||||  
Db 373 TPVAPT 378  
  
RESULT 6  
GATA\_HELPY  
ID GATA\_HELPY STANDARD; PRT; 453 AA.  
AC PS6114:  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE GLUTAMYL-TRNA(GLN) AMIDOTRANSFERASE SUBUNIT A (EC 6.3.5.-) (GLU-ADT  
DE SUBUNIT A).  
GN GATA OR HP0830.  
OS Helicobacter pylori (Campylobacter pylori).  
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;  
OC Helicobacter.  
OX NCBI\_TaxID=210;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=26695 / ATCC 700392;  
RX MEDLINE=9739467; Pubmed=9252185;  
RA Tomb J.-F., White O., Kervavage A.R., Clayton R.A., Sutton G.G.,  
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,  
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
RA Colton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,  
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
RA Venter J.C.;  
RT "The complete genome sequence of the gastric pathogen Helicobacter  
RT pylori".  
RL Nature 388:539-547(1997).  
CC -1- FUNCTION: FURNISHES A MEANS FOR FORMATION OF CORRECTLY CHARGED  
CC GLN-TRNA(GLN) THROUGH THE TRANSMUTATION OF MISACETYLATED GLU-  
CC TRNA(GLN) IN ORGANISMS WHICH LACK GLUTAMINYL-TRNA SYNTHETASE. THE  
CC REACTION TAKES PLACE IN THE PRESENCE OF GLUTAMINE AND ATP THROUGH  
CC AN ACTIVATED GAMMA-PHOSPHO-GLU-TRNA(GLN) (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMYL-TRNA(GLN) -> ADP  
CC + PHOSPHATE + L-GLUTAMINYL-TRNA(GLN) + L-GLUTAMINE.  
CC -1- SUBUNIT: HETEROTRIMER OF A, B AND C SUBUNITS (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>  
CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
CC -----  
CC EMBL: AE000594; AAD07880.1; -  
CC TIGR: HP0830; -  
CC InterPro: IPR000120; -  
CC Pfam: PF01425; AMIDase; 1.  
CC PROSITE: PS00571; AMIDASES; 1.  
CC KEGG: Protein biosynthesis; Ligase.

SO SEQUENCE 453 AA; 49652 MW; A35814B32F1AE13A CRC64;  
  
Query Match  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 TPVAPT 8  
|||||  
Db 373 TPVAPT 378  
  
RESULT 7  
CADA\_LISMO  
ID CADA\_LISMO STANDARD; PRT; 711 AA.  
AC 060048:  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE PROBABLE CADMIUM-TRANSPORTING ATPASE (EC 3.6.1.-) (CADMIUM EFFLUX  
DE ATPASE).  
GN CADA.  
OS Listeria monocytogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/staphylococcus group; Listeria.  
OX NCBI\_TaxID=1639;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LM74;  
RX MEDLINE=94245633; Pubmed=8188605;  
RA Lebrun M., Audurier A., Cossart P.;  
RT "Plasmid-borne cadmium resistance genes in Listeria monocytogenes are  
RT similar to cda and cadC of Staphylococcus aureus and are induced by  
RT cadmium".  
RL J. Bacteriol. 176:3040-3048(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LM74; TRANSPOSON=Trn5422;  
RX MEDLINE=94245634; Pubmed=8188606;  
RA Lebrun M., Audurier A., Cossart P.;  
RT "Plasmid-borne cadmium resistance genes in Listeria monocytogenes are  
RT present on trn5422, a novel transposon closely related to trn917".  
RL J. Bacteriol. 176:3049-3061(1994).  
CC -1- FUNCTION: THIS ELECTRONEUTRAL ANTI-PORTER EFFECTS ONE CD(2+) WHILE  
CC ACCUMULATING TWO PROTONS BY AN ENERGY-DEPENDENT EFFLUX MECHANISM.  
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O -> ADP + ORTHOPHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY  
CC (E1-E2 ATPASES).  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>  
CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
CC -----  
CC EMBL: L28104; AAA25275.1; -  
CC InterPro: IPR001366; -  
CC InterPro: IPR001757; -  
CC InterPro: IPR001934; -  
CC Pfam: PF00122; E1-E2 ATPase; 1.  
CC Pfam: PF00403; HMA; 1.  
CC PRINTS: PR00941; CDATPASE.  
CC PROSITE: PS00154; ATPASE\_E1\_E2; 1.  
CC KEGG: Hydrolyase; Transmembrane; Phosphorylation; ATP-binding; Metal-binding;  
CC Cadmium; Plasmid; Cadmium resistance; Transposable element.  
FT DOMAIN 9 38 HMA.  
FT METAL 14 14 CADMIUM (POTENTIAL).



FT METAL 17 17 CADMIUM (POTENTIAL).  
 FT MOD\_RES 398 398 PHOSPHORYLATION (PROBABLE).  
 SQ SEQUENCE 711 AA: 77089 MW: C23BFEB7E20CEE9A CRC64;

Query Match 66.7%; Score 6; DB 1; Length 711;  
 Best Local Similarity 100.0%; Pred. No. 19;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVA 6  
 |||||  
 Db 360 VSTPVA 365

## RESULT 8

CPN\_DROME STANDARD; PRT: 865 AA.  
 ID CPN\_DROME 002910;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 01-FEB-1994 (Rel. 28, Last annotation update)  
 DE CALPHOTIN.  
 GN CPN OR CAP.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CANTON-S;  
 RX MEDLINE=93165729; PubMed=8094559;  
 RA Martin J.H., Benzer S., Rudnicka M., Miller C.A.;  
 RL "Calphotin: a Drosophila photoreceptor cell calcium-binding protein.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1531-1535(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CANTON-S;  
 RX MEDLINE=93165730; PubMed=8434015;  
 RA Ballinger D.G., Xue N., Harshman K.D.;  
 RL "Calcium and contains a leucine zipper.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:1536-1540(1993).  
 CC -1- FUNCTION: MIGHT FUNCTION AS A CALCIUM-SENSITIZING "SPONGE" TO  
 CC REGULATE THE AMOUNT OF FREE CYTOPLASMIC CALCIUM. IT BINDS 0.3 MOL  
 CC OF CA+2 PER MOL OF PROTEIN.  
 CC -1- SUBUNIT: HOMODIMER (PROBABLE).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC; HYPODERMIS COMPARTMENT.  
 CC -1- TISSUE SPECIFICITY: SOMA AND AXONS OF PHOTORECEPTOR CELLS OF  
 CC COMPOUND EYES AND OCELLI.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN PHOTORECEPTOR CELL  
 CC DEVELOPMENT.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: L02111; AAA28405.1; -;  
 CC EMBL: L05080; AAA28420.1; -;  
 DR PIR: A47282; A47282.  
 DR Flybase: FBgn010218; Cpn.  
 KM Calcium-binding.  
 FT CONFLICT 36 36 A -> AVAPAVA (IN REF. 2).  
 FT CONFLICT 43 43 I -> T (IN REF. 2).  
 FT CONFLICT 64 64 I -> V (IN REF. 2).  
 FT CONFLICT 76 76 T -> A (IN REF. 2).  
 FT CONFLICT 100 100 P -> P (IN REF. 2).  
 FT CONFLICT 126 127 VO -> AP (IN REF. 2).  
 FT CONFLICT 154 154 I -> V (IN REF. 2).

FT CONFLICT 160 160 S -> T (IN REF. 2).  
 FT CONFLICT 534 534 A -> E (IN REF. 2).  
 FT CONFLICT 699 699 I -> T (IN REF. 2).  
 FT CONFLICT 703 703 V -> L (IN REF. 2).  
 FT CONFLICT 721 721 D -> E (IN REF. 2).  
 SQ SEQUENCE 865 AA: 84781 MW: 2110417E0B0E7CFE CRC64;

Query Match 66.7%; Score 6; DB 1; Length 865;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVA 6  
 |||||  
 Db 97 VSTPVA 102

## RESULT 9

RM1\_DROME STANDARD; PRT: 1540 AA.  
 ID RM1\_DROME 09V7H4;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE RM1 PROTEIN HOMOLOG.  
 GN CG8370.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borzova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Daventport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.D., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hovis L., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jajala B.E., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.F., Smith T.,  
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtikas R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.-A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AEO03808; AAF58081.1; -  
DR FlyBase: FBgn0034060; CG8370.  
SO SEQUENCE 1540 AA; 170274 MW; 5AB097531D425846 CRC64;  
  
Query Match 66.7%; Score 6; DB 1; Length 1540;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 VSTPVA 6  
| | | | |  
DB 1207 VSTPVA 1212  
  
RESULT 10  
KAC4\_RABIT STANDARD; PRT; 103 AA.  
ID KAC4\_RABIT  
AC P01840;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE IG KAPPA-B4 CHAIN C REGION.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83300036; PubMed=6412231;  
RA Emorine L., Dreher K.L., Kindt T.J., Max E.E.;  
RT "Rabbit immunoglobulin kappa genes: structure of a germline b4  
RT allotype J-C locus and evidence for several b4-related sequences in  
RT the rabbit genome.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:5709-5713(1983).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=82060334; PubMed=6795636;  
RA Heidmann O., Auffray C., Cazenave P.-A., Rougeon F.;  
RT "Nucleotide sequence of constant and 3' untranslated regions of a  
RT kappa immunoglobulin light chain mRNA of a homozygous b4 rabbit.";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:5802-5806(1981).  
RN [3]  
RP SEQUENCE.  
RX MEDLINE=75133568; PubMed=1091650;  
RA Chen K.C.S., Kindt T.J., Krause R.M.;  
RT "Primary structure of the L chain from a rabbit homogeneous antibody  
RT to streptococcal carbohydrate. II. Sequence determination of peptides  
RT from tryptic and peptic digests.";  
RL J. Biol. Chem. 250:3289-3296(1975).  
CC -1- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM ANTIBODY TO THE  
CC SPECIFIC CARBOHYDRATE OF GROUP C STREPTOCOCCI & WAS ISOLATED FROM  
CC THE SERUM OF A SINGLE RABBIT.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X00231; CAA25051.1; -  
DR PIR: A02122; KARB.  
DR InterPro: IPR000495; -

-----  
DR InterPro: IPR003006; -  
DR Pfam: PF00047; 1g. 1.  
DR PROSITE: PS00290; IG\_MHC; FALSE NEG.  
KW Immunoglobulin domain; Immunoglobulin C region.  
FT NON\_TER 1  
FT DISULFID 26 85  
FT DISULFID 103 103 INTERCHAIN (WITH A HEAVY CHAIN).  
FT CONFLICT 58 58 N -> D (IN REF. 3).  
SO SEQUENCE 103 AA; 11043 MW; 5FC5AC8B60E68DB CRC64;  
  
Query Match 55.68%; Score 5; DB 1; Length 103;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 4 PVAPT 8  
| | | | |  
DB 2 PVAPT 6  
  
RESULT 11  
SY12\_MOUSE STANDARD; PRT; 104 AA.  
ID SY12\_MOUSE  
AC Q62401;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE SMALL INDUCIBLE CYTOKINE A12 PRECURSOR (MONOCYTE CHEMOTACTIC PROTEIN  
DE 5) (MCP-5) (MCP-1 RELATED CHEMOKINE).  
GN SCYA12 OR MCP5.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97079149; PubMed=8920881;  
RA Jia G.-Q., Gonzalo J.A., Lloyd C., Kremer L., Lu L., Martinez A.C.,  
RA Wershl B.K., Gutierrez-Ramos J.C.;  
RT "Distinct expression and function of the novel mouse chemokine  
RT monocyte chemoattractant protein-5 in lung allergic inflammation.";  
RL J. Exp. Med. 184:1939-1951(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97149438; PubMed=8996246;  
RA Sarraf M.N., Garcia-Zepeda E.A., Molean J.A., Charo I.F., Luster A.D.;  
RT "Murine monocyte chemoattractant protein (MCP)-5: a novel CC  
RT chemokine that is a structural and functional homologue of human  
RT MCP-1.";  
RL J. Exp. Med. 185:99-109(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX SPRAT-B10.S/J, BALB/C, DBA/2J, AND NOD/LTJ; TISSUE=Spleen;  
RX MEDLINE=99370037; PubMed=10438970;  
RA Tauscher C., Butterfield R.J., Ma R.Z., Zachary J.F., Doerge R.W.,  
RA Blankenhorn E.P.;  
RT "Sequence polymorphisms in the chemokines Sycal (TCA-3), Sycal2  
RT (monocyte chemoattractant protein (MCP)-1), and Sycal2 (MCP-5) are  
RT candidates for eae7, a locus controlling susceptibility to monophasic  
RT remitting/nonrelapsing experimental allergic encephalomyelitis.";  
RL J. Immunol. 163:2262-2266(1999).  
CC -1- FUNCTION: CHEMOTACTIC FACTOR THAT ATTRACTS EOSINOPHILS, MONOCYTES,  
CC AND LYMPHOCYTES BUT NOT NEUTROPHILS. POTENT MONOCYTE ACTIVE  
CC CHEMOKINE THAT SIGNALS THROUGH CCR2. INVOLVED IN ALLERGIC  
CC INFLAMMATION AND THE HOST RESPONSE TO PATHOGENS AND MAY PLAY A  
CC PIVOTAL ROLE DURING EARLY STAGES OF ALLERGIC LUNG INFLAMMATION.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE LYMPH NODES AND  
CC THYMUS. ALSO FOUND IN THE SALIVARY GLANDS CONTAINING LYMPH NODES,  
CC BREAST, HEART, LUNG, BRAIN, SMALL INTESTINE, KIDNEY AND COLON.  
CC -1- INDUCTION: BY IFN-GAMMA AND LIPOPOLYSACCHARIDE (LPS).  
CC -1- SIMILARITY: BELONGS TO THE INTERCRINE BETA FAMILY (SMALL CYTOKINE

CC (C-C) (CHEMOKINE CC).  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: U50712; AAB50053.1; -;  
 DR EMBL: U66670; AAB49424.1; -;  
 DR EMBL: AF065934; AAF15384.1; -;  
 DR EMBL: AF065935; AAF15385.1; -;  
 DR EMBL: AF065936; AAF15386.1; -;  
 DR EMBL: AF065937; AAF15387.1; -;  
 DR HSSP: P13500; IDOL.  
 DR MGD: MGI:108224; Scya12.  
 DR InterPro: IPR000827; -;  
 DR InterPro: IPR001811; -;  
 DR Pfam: PF00048; I18; 1.  
 DR PROSITE: PS00472; SMALL\_CYTOKINES\_CC; 1.  
 DR Cytokine; Chemotaxis; Signal; Inflammatory response.  
 KW SIGNAL 1 22 BY SIMILARITY.  
 FT CHAIN 23 104 SMALL INDUCIBLE CYTOKINE A12.  
 FT DISULFID 33 58 BY SIMILARITY.  
 FT DISULFID 34 74 BY SIMILARITY.  
 SQ SEQUENCE 104 AA: 11659 MW: 80102P4FF4CC3DBF CRC64:  
 -----  
 Query Match 55.6%; Score 5; DB 1; Length 104;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 -----  
 Oy 1 VSTPV 5  
 11111  
 Db 27 VSTPV 31  
 -----  
 RESULT 12  
 PSSP\_HUMAN STANDARD; PRT; 114 AA.  
 ID PSSP\_HUMAN P08118; P11999;  
 AC 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PROSTATE SECRETED SEMINAL PLASMA PROTEIN PRECURSOR (PSP-94) (BETA-  
 DE MICROSEMINOPROTEIN) (SEMINAL PLASMA BETA-INHIBIN) (IMMUNOGLOBULIN  
 DE BINDING FACTOR) (IGBF) (PROSTATE SECRETORY PROTEIN PSP94) (FN44).  
 GN MSMB OR PSP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87161231; PubMed=3829888;  
 RA Mdkay M., Nolet S., Fournier S., Benjannet S., Chapdelaine P.,  
 RA Mdkay G., Dube J.Y., Tremblay R., Lazure C., Seidah N.G.,  
 RA Chretien M.;  
 RA "Molecular cloning and sequence of the cDNA for a 94-amino-acid  
 RT seminal plasma protein secreted by the human prostate.";  
 RL DNA 6:23-29(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91274357; PubMed=2054385;  
 RA Nolet S., Mdkay M., Chretien M.;  
 RT "Prostatic secretory protein PSP94: gene organization and promoter  
 RT sequence in rhesus monkey and human.";  
 RL Biochim. Biophys. Acta 1089:247-249(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90211299; PubMed=2322265;  
 -----

RA Green C.B., Liu W.Y., Kwok S.C.M.;  
 RT "Cloning and nucleotide sequence analysis of the human beta-  
 RT microseminoprotein gene.";  
 RL Biochem. Biophys. Res. Commun. 167:1184-1190(1990).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90073664; PubMed=2590204;  
 RA Ulyshack M., Lindstroem C., Welber H., Abrahamsson P.-A., Lilla H.,  
 RA Lundvall A.;  
 RT "Molecular cloning of a small prostate protein, known as beta-  
 RT microseminoprotein, PSP94 or beta-inhibin, and demonstration of  
 RT transcripts in non-genital tissues.";  
 RL Biochem. Biophys. Res. Commun. 164:1310-1315(1989).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Prostate;  
 RX MEDLINE=94115955; PubMed=7506990;  
 RA Liu A.Y., Bradner R.C., Vessella R.L.;  
 RT "Decreased expression of prostatic secretory protein PSP94 in  
 RT prostate cancer.";  
 RL Cancer Lett. 74:91-99(1993).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Bajjal-Gupta M., Clarke M.W.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE OF 21-113.  
 RX MEDLINE=8519974; PubMed=3995056;  
 RA Akiyama K., Yoshioke Y., Schmid K., Offner G.D., Troxler R.F.,  
 RA Tsuda R., Hara M.;  
 RT "The amino acid sequence of human beta-microseminoprotein.";  
 RL Biochim. Biophys. Acta 829:288-294(1985).  
 RN [8]  
 RP SEQUENCE OF 21-114.  
 RX MEDLINE=85004133; PubMed=6434350;  
 RA Seidah N.G., Arbatli N.J., Rochemont J., Sheth A.R., Chretien M.;  
 RT "Complete amino acid sequence of human seminal plasma beta-inhibin.  
 RT Prediction of post Glu-Arg cleavage as a maturation site.";  
 RL FEBS Lett. 175:349-355(1984).  
 RN [9]  
 RP SEQUENCE OF 21-50 AND 113-114.  
 RX MEDLINE=92028964; PubMed=1930232;  
 RA Liang Z.G., Kamada M., Koide S.S.;  
 RT "Structural identity of immunoglobulin binding factor and prostatic  
 RT secretory protein of human seminal plasma.";  
 RL Biochem. Biophys. Res. Commun. 180:356-359(1991).  
 CC -1- FUNCTION: INHIBITS THE SECRETION OF FSH BY PITUITARY CELLS.  
 CC -1- SUBCELLULAR LOCATION: SPERM SURFACE.  
 CC -1- PTM: RICH IN DISULFIDE BONDS (POTENTIAL).  
 CC -1- MISCELLANEOUS: SPECIFIC RECEPTORS FOR THIS PROTEIN ARE FOUND ON  
 CC SPERMATOZOA AND IN THE PROSTATE.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: M34376; AAB59871.1; -;  
 DR EMBL: M34373; AAB59871.1; JOINED.  
 DR EMBL: M34374; AAB59871.1; JOINED.  
 DR EMBL: M34375; AAB59871.1; JOINED.  
 DR EMBL: M15885; AAB59871.1; -;  
 DR EMBL: M57928; CAA41002.1; -;  
 DR EMBL: X57929; CAA41002.1; JOINED.  
 DR EMBL: X57930; CAA41002.1; JOINED.  
 DR EMBL: X57931; CAA41002.1; JOINED.  
 DR EMBL: S67815; AAB29732.1; -;  
 DR EMBL: U78976; AAB37355.1; -;  
 DR PIR: A26451.  
 DR PIR: A29777; A29777.  
 -----

DR PIR: A30984; A30984.  
 DR PIR: A34567; A34567.  
 DR PIR: S16238; S16238.  
 DR MIM: 157145; -.  
 KM Sperm: Signal.  
 FT SIGNAL 1 20  
 FT CHAIN 21 114  
 FT CONFLICT 59 60  
 FT CONFLICT 113 113  
 FT CONFLICT 113 113  
 SO SEQUENCE 114 AA; 12865 MW; DOB13E436D70B8ID CRC64;  
 TC -> PT (IN REF. 7).  
 I -> G (IN REF. 8).  
 Query Match 55.6%; Score 5; DB 1; Length 114;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 VSTPV 5  
 Db 73 VSTPV 77  
 RESULT 13  
 YELL\_SCHPO STANDARD; PRT; 166 AA.  
 AC 013883;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE HYPOTHETICAL PROTEIN CIB3.19 IN CHROMOSOME I (FRAGMENT).  
 GN SPACIB3.19.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z98598; CAB1245.1; -.  
 KM Hypothetical protein; Transmembrane.  
 FT TRANSMEM 9 29  
 FT TRANSMEM 66 86  
 FT TRANSMEM 88 108  
 FT TRANSMEM 125 145  
 FT NON\_TER 166 166  
 FT NON\_TER 166 166  
 SO SEQUENCE 166 AA; 18657 MW; 1B4A67978321D667 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 VSTPV 5  
 Db 38 VSTPV 42  
 RESULT 14  
 CASK\_RAT STANDARD; PRT; 178 AA.  
 ID CASK\_RAT  
 AC P04466;

DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE KAPPA CASEIN PRECURSOR.  
 GN CSN10 OR CSN3 OR CSNK.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85054984; PubMed=6094580;  
 RA Nekhai H.L., Granham F.H., Guillino P.M.;  
 RT "Expression of kappa-casein in normal and neoplastic rat mammary  
 RT gland is under the control of prolactin."  
 RL J. Biol. Chem. 259:14894-14898(1984).  
 CC -1- FUNCTION: KAPPA CASEIN STABILIZES MICELLE FORMATION, PREVENTING  
 CC CASEIN PRECIPITATION IN MILK.  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -1- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.  
 CC -1- SIMILARITY: BELONGS TO THE KAPPA-CASEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; K02598; AAA40880.1; -.  
 DR PIR: A03115; KKRT.  
 DR Interpro: IPR000117; -.  
 DR Pfam: PF00997; casein.kappa.1.  
 KM Milk; Glycoprotein; phosphorylation; signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 178  
 FT SITE 115 116  
 FT MOD\_RES 120 120  
 FT MOD\_RES 145 145  
 FT MOD\_RES 165 165  
 SO SEQUENCE 178 AA; 19548 MW; 0BC793AEF8D7423 CRC64;  
 CLEAVAGE (BY CHYMOSIN (=RENNIN)).  
 PHOSPHORYLATION (POTENTIAL).  
 PHOSPHORYLATION (POTENTIAL).  
 PHOSPHORYLATION (POTENTIAL).

Query Match 55.6%; Score 5; DB 1; Length 178;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 VSTPV 5  
 Db 139 VSTPV 143  
 RESULT 15  
 YFL5\_YEAST STANDARD; PRT; 178 AA.  
 ID YFL5\_YEAST  
 AC P43617;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE PUTATIVE MITOCHONDRIAL CARRIER YFR045W.  
 GN YFR045W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=S288C / AB972;  
 RX MEDLINE=95400292; PubMed=7670463;  
 RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
 RA Sasamura S.-I., Sasamura M., Tsuchiya Y., Soeda E., Yokoyama K.,  
 RA Yamazaki M., Tashiro H., Eki T.;

```

RT "Analysis of the nucleotide sequence of chromosome VI from
RL Saccharomyces cerevisiae."
RN Nat. Genet. 10:261-268(1995).
RP [2]
RC SEQUENCE FROM N.A.
RA STRAIN-S288C / AB972;
RX MEDLINE=96287652; PubMed=8686379;
RA Ekl T., Naitou M., Hagihara H., Ozawa M., Sasamura S.-I.,
RA Sasamura M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;
RT "Analysis of a 36.2 kb DNA sequence including the right telomere of
RT chromosome VI from Saccharomyces cerevisiae."
RL yeast 12:149-167(1996).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN, MITOCHONDRIAL
CC INNER MEMBRANE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: D50617; BAA09284.1; -
DR EMBL: D44597; BAA08011.1; -
DR SGD: S0001941; YF045W.
DR InterPro: IPR001993; -
DR Pfam: PF00153; MITOCH_CARR; 1.
DR PROSITE: PS00215; MITOCH_CARRIER; 1.
KW Hypothetical protein; Mitochondrion; Inner membrane; Repeat;
KW Transmembrane; Transport.
SQ SEQUENCE 178 AA; 19901 MW; D57433A7D1AA4FE5 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 178;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 STPVA 6
| | | | |
DB 27 STPVA 31

RESULT 16
ATPD_BACP3 STANDARD; PRT; 179 AA.
AC P09220;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ATP SYNTHASE DELTA CHAIN (EC 3.6.1.34);
GN ATPH.
OS Bacillus PS3 (Thermophilic bacterium PS-3).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=70306;
RN [1]
RC SEQUENCE FROM N.A.
RA MEDLINE=88163679; PubMed=2894854;
RA Ohta S., Yonda M., Ishizuka M., Hirata H., Hamamoto T.,
RA Ohtawa-Hamamoto Y., Matsuda K., Kagawa Y.;
RT "Sequence and over-expression of subunits of adenosine triphosphate
RT synthase in thermophilic bacterium PS3."
RL Biochim. Biophys. Acta 933:141-155(1988).
CC -1- FUNCTION: THIS PROTEIN SEEMS TO BE PART OF THE STALK THAT LINKS
CC CF(0) TO CF(1). IT EITHER TRANSITS CONFORMATIONAL CHANGES FROM
CC CF(0) INTO CF(1) OR IS IMPLICATED IN PROTON CONDUCTION.
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- SIMILARITY: BELONGS TO THE ATPASE DELTA CHAIN FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: X07804; CAA30651.1; -
DR PIR: S01400; S01400.
DR InterPro: IPR000711; -
DR Pfam: PF00213; OSCP; 1.
DR PROSITE: PS00389; ATPASE_DELTA; 1.
KW Hydrolyase; ATP synthesis; CF(1); Hydrogen ion transport.
SQ SEQUENCE 179 AA; 19657 MW; 8168E151121380E5 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 179;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTPV 5
| | | | |
DB 69 VSTPV 73

RESULT 17
CASK_MOUSE STANDARD; PRT; 181 AA.
AC P06796;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KAPPA CASEIN PRECURSOR.
GN CSN10 OR CSN3 OR CSNK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RA MEDLINE=66004062; PubMed=4042811;
RA Thompson M.D., Dave J.R., Nakhasi H.L.;
RT "Molecular cloning of mouse mammary gland kappa-casein: comparison
RT with rat kappa-casein and rat and human gamma-fibrinogen."
RL DNA 4:263-271(1985).
CC -1- FUNCTION: KAPPA CASEIN STABILIZES MICELLE FORMATION, PREVENTING
CC CASEIN PRECIPITATION IN MILK.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -1- SIMILARITY: BELONGS TO THE KAPPA-CASEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: M10114; AAA37370.1; -
DR PIR: A05080; A05080.
DR MGD: MGI:107461; Csnk.
DR InterPro: IPR000117; -
DR Pfam: PF00997; casein_kappa; 1.
KW Milk; Glycoprotein; Phosphorylation; Signal.
FT SIGNAL 1 21
FT CHAIN 22 181 KAPPA CASEIN.
FT SITE 118 119 CLEAVAGE (BY CHYMOSIN (=RENNIN)).
FT MOD_RES 123 123 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 181 AA; 20062 MW; E5ABEF01679C5DC CRC64;

```

Query Match 55.6%; Score 5; DB 1; Length 181;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 |||||  
 Db 142 VSTPV 146

RESULT 18  
 V33P\_ADE41 STANDARD; PRT; 217 AA.  
 AC P19416; Q64818;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 01-MAR-1992 (Rel. 21, Last annotation update)  
 DE 33 KDa PHOSPHOPROTEIN.  
 OS Human adenovirus type 41.  
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.  
 NCBI\_TaxID=10524;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-TAK;  
 RC MEDLINE=90272433; PubMed=2349115;  
 RA Stemenda S.B., Pieniazek N.J., Velarde J. Jr., Pieniazek D.,  
 Luftig R.B.;  
 RT "Nucleotide sequence of the region coding for 100K and 33K proteins  
 of human enteric adenovirus type 41 (Tak).";  
 RL Nucleic Acids Res. 18:3069-3069(1990).  
 RN [2]  
 RN SEQUENCE OF 203-217 FROM N.A.  
 RC STRAIN-TAK;  
 RX MEDLINE=89345113; PubMed=2762136;  
 RA Pieniazek N.J., Velarde J. Jr., Pieniazek D., Luftig R.B.;  
 RT "Nucleotide sequence of human enteric adenovirus type 41 hexon-  
 associated protein VIII precursor (pVIII) including the early region  
 E3 promoter.";  
 RL Nucleic Acids Res. 17:5398-5398(1989).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X52532; CAB38632.1; -;  
 DR EMBL: X15137; CAA33236.1; -;  
 DR PIR: S04851; S04851.  
 DR PIR: S10212; S10212.  
 KW Late protein; Phosphorylation.  
 SQ SEQUENCE 217 AA; 24740 MW; B6ECDAB3C24EC3E CRC64;

Query Match 55.6%; Score 5; DB 1; Length 217;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
 |||||  
 Db 118 PVAPT 122

RESULT 19  
 RPE\_TREPA STANDARD; PRT; 218 AA.  
 ID RPE\_TREPA  
 AC 066107;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE RIBULOSE-PHOSPHATE 3-EPIMERASE (EC 5.1.3.1) (PENTOSE-5-PHOSPHATE 3-

DE EPIMERASE) (PPE) (R5P3E).  
 GN RPE OR TP0945.  
 OS Treponema pallidum.  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
 OX NCBI\_TaxID=160;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP Shevchenko D.V., Akins D.R., Radolf J.D.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN-NICHOLS;  
 RX MEDLINE=98332770; PubMed=9665876;  
 RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,  
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,  
 RA Sdergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uitterback T.,  
 RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
 RA Venter J.C.;  
 RT "Complete genome sequence of Treponema pallidum, the syphilis  
 agent."  
 RL Science 281:375-388(1998).  
 CC -I- CATALYTIC ACTIVITY: D-RIBULOSE 5-PHOSPHATE - D-XYLULOSE 5-PHOSPHATE.  
 CC -I- SIMILARITY: BELONGS TO THE RIBULOSE-PHOSPHATE 3-EPIMERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U97573; AAC08057.1; -;  
 DR EMBL: AE001263; AAC65902.1; -;  
 DR TIGR: TP0945; -;  
 DR InterPro: IPR000056; -;  
 DR Pfam: PF00834; Ribul\_P\_3\_epim; 1.  
 DR PROSITE: PS01085; RIBUL\_P\_3\_EPIMER\_1; 1.  
 DR PROSITE: PS01086; RIBUL\_P\_3\_EPIMER\_2; 1.  
 KW Isomerase; Carbohydrate metabolism.  
 SQ SEQUENCE 218 AA; 23514 MW; 0066263301FA2PCA CRC64;

Query Match 55.6%; Score 5; DB 1; Length 218;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
 |||||  
 Db 120 STPVA 124

RESULT 20  
 Y574\_AERPE STANDARD; PRT; 225 AA.  
 ID Y574\_AERPE  
 AC Q9YK4;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE PUTATIVE SHIKIMATE 5-DEHYDROGENASE-LIKE PROTEIN APE0574.  
 GN APE0574.  
 OS Aeropyrum pernix.  
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;  
 OC Aeropyrum.  
 OX NCBI\_TaxID=56636;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-K1;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,

RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankael A., Kosugi H.,  
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,  
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,  
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,  
RT "Complete genome sequence of an aerobic hyper-thermophilic  
RT crenarchaeon, Aeropyrum pernix K1.";  
RL DNA Res. 6:83-101(1999).  
CC -1- SIMILARITY: BELONGS TO THE SHIKIMATE DEHYDROGENASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AP000060; BAA79542.1; -  
DR Hypothetical protein: Oxidoreductase.  
SO SEQUENCE 225 AA; 23273 MW; 13110B4615003F6C CRC64;

Query Match 55.6%; Score 5; DB 1; Length 225;  
Best Local Similarity 100.0%; Pred. No. 79;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
| | | | |  
Db 147 VSTPV 151

RESULT 21  
PRL\_BOVIN STANDARD: PRT: 229 AA.  
ID P01239; Q95112; Q29417;  
AC 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PROLACTIN PRECURSOR (PRL).  
GN PRL.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=82098114; PubMed=6274859;  
RA Sasavage N.L., Nilson J.H., Horowitz S., Rottman F.M.;  
RT "Nucleotide sequence of bovine prolactin messenger RNA. Evidence for  
RT sequence polymorphism.";  
RL J. Biol. Chem. 257:678-681(1982).  
RN [2]  
RP SEQUENCE OF 21-229 FROM N.A.  
RX MEDLINE=83157107; PubMed=6299665;  
RA Miller W.L., Colt D., Baxter J.D., Martial J.A.;  
RT "Cloning of bovine prolactin cDNA and evolutionary implications of  
RT its sequence.";  
RL DNA 1:37-50(1981).  
RN [3]  
RP REVISIONS.  
RX MEDLINE=83182016; PubMed=6897772;  
RA Miller W.L., Colt D., Baxter J.D., Martial J.A.;  
RT "Bovine prolactin: corrected cDNA sequence and genetic  
RT polymorphisms.";  
RL DNA 1:313-314(1982).  
RN [4]  
RP SEQUENCE OF 96-229 FROM N.A.  
RX Rubtsov P.M., Oganesyan R.G., Gorbulev V.G., Skryabin K.G., Baev A.A.;  
RT "Genetic engineering of peptide hormones. II. Possible polymorphism of  
RT preprolactin in cattle. Data of molecular cloning.";  
RL Mol. Biol. (Mosk) 22:117-127(1988).

RN [5]  
RP PRELIMINARY SEQUENCE OF 31-229.  
RX MEDLINE=75031394; PubMed=4608931;  
RA Wallis M.;  
RT "The primary structure of bovine prolactin.";  
RL FEBS Lett. 44:205-208(1974).  
RN [6]  
RP SEQUENCE OF 31-46.  
RX MEDLINE=71150631; PubMed=5507606;  
RA Graf L., Cseh G., Nagy I., Kurcz M.;  
RT "An evidence for deamidation of prolactin monomer.";  
RL Acta Biochim. Biophys. Acad. Sci. Hung. 5:299-303(1970).  
RN [7]  
RP PHOSPHORYLATION SITES.  
RX TISSUE-Pituitary;  
RC MEDLINE=94071839; PubMed=8250856;  
RA Kim B.G., Brooks C.L.;  
RT "Isolation and characterization of phosphorylated bovine prolactin.";  
RL Biochem. J. 296:41-47(1993).  
CC -1- FUNCTION: PROLACTIN ACTS PRIMARILY ON THE MAMMARY GLAND BY  
CC PROMOTING LACTATION.  
CC -----  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: V00112; CAA23446.1; -  
DR EMBL: X01452; -; NOT\_ANNOTATED\_CDS.  
DR EMBL: X01744; CAA25880.1; -  
DR EMBL: M36873; AAA30737.1; -  
DR EMBL: M36874; AAA30738.1; -  
DR EMBL: X14320; CAA32500.1; -  
DR EMBL: X14321; CAA32501.1; -  
DR PIR: A01508; LCBO.  
DR HSSP: Q28632; IAN3.  
DR InterPro: IPR001400; -  
DR Pfam: PF00103; hormone.1.  
DR PRINTS: PR00836; SOMATOTROPIN.  
DR PROSITE: PS00266; SOMATOTROPIN\_1; 1.  
DR PROSITE: PS00338; SOMATOTROPIN\_2; 1.  
KW Hormone; Parturition; Lactation; Pituitary; Signal; Phosphorylation.  
FT CHAIN 1 30  
FT SIGNAL 1 30  
FT DISULFID 34 41  
FT DISULFID 88 204  
FT DISULFID 221 229  
FT MOD\_RES 56 56 PHOSPHORYLATION.  
FT MOD\_RES 64 64 PHOSPHORYLATION.  
FT MOD\_RES 120 120 PHOSPHORYLATION.  
FT CONFLICT 61 61 D -> N (IN REF. 5).  
SO SEQUENCE 229 AA; 25792 MW; E7E9BB655A26F3D CRC64;

Query Match 55.6%; Score 5; DB 1; Length 229;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
| | | | |  
Db 29 VSTPV 33

RESULT 22  
PRL\_CAPHI STANDARD: PRT: 229 AA.  
ID Q28318; Q28329;  
AC 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, last sequence update)  
 DT 01-NOV-1997 (Rel. 35, last annotation update)  
 DE PROLACTIN PRECURSOR (PRL).  
 GN PRL.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95059806; PubMed=7969789;  
 RA Le Provost F., Leroux C., Martin P., Gaye P., Djiane J.;  
 RL "Prolactin gene expression in ovine and caprine mammary gland.";  
 CC Neuroendocrinology 60:305-313(1994).  
 CC -1- FUNCTION: PROLACTIN ACTS PRIMARILY ON THE MAMMARY GLAND BY  
 PROMOTING LACTATION.  
 CC -1- SUBCELLULAR LOCATION: SECRETED  
 CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X76049; CAA53634.1; -;  
 DR EMBL: X76048; CAA53633.1; -;  
 DR HSSP: Q28632; IAN3.  
 DR InterPro: IPR001400; -;  
 DR Pfam: PF00103; hormone; 1.  
 DR PRINTS: PR00836; SOMATOTROPIN.  
 DR PROSITE: PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE: PS00338; SOMATOTROPIN\_2; 1.  
 DR Hormone; Parturition; Lactation; Pituitary; Signal.  
 KW hormone; Parturition; Lactation; Pituitary; Signal.  
 FT CHAIN 1 30 BY SIMILARITY.  
 FT DISULFID 31 229 BY SIMILARITY.  
 FT DISULFID 34 41 BY SIMILARITY.  
 FT DISULFID 88 204 BY SIMILARITY.  
 FT DISULFID 221 229 BY SIMILARITY.  
 FT VARIANT 164 164 L->F.  
 FT VARIANT 164 164 L->F.  
 SQ SEQUENCE 229 AA; 25773 MW; 331C640C66134DDA CRC64;

Query Match 55.6%; Score 5; DB 1; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 Db 29 VSTPV 33

RESULT 23  
 PRL\_SHEEP STANDARD; PRT; 229 AA.  
 AC P01240; Q28687;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE PROLACTIN PRECURSOR (PRL).  
 GN PRL.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pituitary;  
 RX MEDLINE=89098399; PubMed=2911473;

RA Adams T.E., Baker L., Brandon M.R.;  
 RT "Cloning and nucleotide sequence of an ovine prolactin cDNA.";  
 RL Nucleic Acids Res. 17:440-440(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pituitary;  
 RX MEDLINE=89326152; PubMed=2666265;  
 RA Varma S., Kwok S., Ebner K.E.;  
 RT "Cloning and nucleotide sequence of ovine prolactin cDNA.";  
 RL Gene 77:349-359(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95059806; PubMed=7969789;  
 RA Le Provost F., Leroux C., Martin P., Gaye P., Djiane J.;  
 RT "Prolactin gene expression in ovine and caprine mammary gland.";  
 RL Neuroendocrinology 60:305-313(1994).  
 RN [4]  
 RP SEQUENCE OF 31-229.  
 RX MEDLINE=71091978; PubMed=5497153;  
 RA Li C.H., Dixon J.S., Lo T.-B., Schmidt K.D., Pankov Y.A.;  
 RT "Studies on pituitary lactogenic hormone. XXX. The primary structure  
 RT of the sheep hormone.";  
 RL Arch. Biochem. Biophys. 141:705-737(1970).  
 RN [5]  
 RP REVISIONS.  
 RX MEDLINE=76189476; PubMed=1270193;  
 RA Li C.H.;  
 RT "Studies on pituitary lactogenic hormone. The primary structure of  
 RT the porcine hormone.";  
 RL Int. J. Pept. Protein Res. 8:205-224(1976).  
 CC -1- FUNCTION: PROLACTIN ACTS PRIMARILY ON THE MAMMARY GLAND BY  
 PROMOTING LACTATION, MAMMOGENESIS, MITOGENESIS AND OSMOREGULATION.  
 CC -1- SUBCELLULAR LOCATION: SECRETED.  
 CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X13483; CAA31839.1; -;  
 DR EMBL: M27057; AAA31578.1; ALT\_INIT.  
 DR EMBL: X76050; CAA53635.1; -;  
 DR PIR: A01509; ICSH.  
 DR PIR: S02104; S02104.  
 DR PIR: J50200; J50200.  
 DR HSSP: Q28632; IAN3.  
 DR InterPro: IPR001400; -;  
 DR Pfam: PF00103; hormone; 1.  
 DR PRINTS: PR00836; SOMATOTROPIN.  
 DR PROSITE: PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE: PS00338; SOMATOTROPIN\_2; 1.  
 DR Hormone; Parturition; Lactation; Pituitary; Signal; Glycoprotein.  
 KW hormone; Parturition; Lactation; Pituitary; Signal; Glycoprotein.  
 FT SIGNAL 1 30  
 FT CHAIN 1 229 PROLACTIN.  
 FT DISULFID 34 41  
 FT DISULFID 88 204  
 FT DISULFID 221 229  
 FT CARBOHYD 61 61  
 FT CONFLICT 40 40 N->D (IN REF. 2).  
 SQ SEQUENCE 229 AA; 25777 MW; 97FD8AF2991B9B39 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 80;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 Db 29 VSTPV 33



RESULT 24  
 HB21\_HUMAN STANDARD: PRT; 261 AA.  
 ID HB21\_HUMAN P01918; P01917;  
 AC P01918; P01917;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DQ(1) BETA CHAIN PRECURSOR  
 (DC-3 BETA CHAIN).  
 GN HLA-DQB.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84298108; PubMed=6206493;  
 RA Boss J.M., Strominger J.L.;  
 RT "Cloning and sequence analysis of the human major histocompatibility  
 complex gene DC-3 beta.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5199-5203(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84031733; PubMed=6415003;  
 RA Larhammar D., Andersson G., Andersson M., Bill P., Boehme J.,  
 Claesson L., Denaro M., Emmoth E., Gustafsson K., Hammarling U.,  
 Heidin E., Hyldig-Nielsen J.J., Lind P., Schemming L., Servenius B.,  
 Widmark E., Rask L., Peterson P.A.;  
 RT "Molecular analysis of human class II transplantation antigens and  
 their genes.";  
 RL Hum. Immunol. 8:95-103(1983).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=82247917; PubMed=6954511;  
 RA Larhammar D., Schemming L., Gustafsson K., Wiman K., Claesson L.,  
 Rask L., Peterson P.A.;  
 RT "Complete amino acid sequence of an HLA-DR antigen-like beta chain as  
 predicted from the nucleotide sequence: similarities with  
 immunoglobulins and HLA-A, -B, and -C antigens.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 79:3687-3691(1982).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97419171; PubMed=9271631;  
 RA Ellis M.C., Heutsma A.H., Ruddy D.A., Hansen S.L., Kronmal G.S.,  
 McClelland E., Quintana L., Drayna D.T., Aldrich M.S., Mignot E.;  
 RT "HLA class II haplotype and sequence analysis support a role for DQ in  
 narcolepsy.";  
 RL Immunogenetics 46:410-417(1997).  
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN WAS PREVIOUSLY CALLED AN HLA-DR  
 BETA CHAIN BY REF.3 AUTHORS.  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC CC  
 CC EMBL: K02405; AAA75521.1; -;  
 CC EMBL: U92032; AAB91991.1; -;  
 CC PIR: A02231; HLHDB.  
 CC HSSP: P13760; 2SERB.  
 CC InterPro: IPR000353; -;  
 CC InterPro: IPR000495; -;  
 CC InterPro: IPR003006; -;  
 CC Pfam: PF00069; MHC\_II\_beta; 1.  
 CC Pfam: PF00047; Ig; 1.  
 CC DR PROSITE: PS00290; IG\_MHC; 1.  
 CC MHC II; Transmembrane; Glycoprotein; signal.  
 FT SIGNAL 1 32

FT CHAIN 33 261 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN,  
 FT DOMAIN 33 126 DQ(1) BETA CHAIN.  
 FT DOMAIN 127 220 EXTRACELLULAR BETA-1.  
 FT DOMAIN 221 230 CONNECTING PEPTIDE.  
 FT TRANSMEM 231 251 CYTOPLASMIC TAIL.  
 FT DOMAIN 252 261 BY SIMILARITY.  
 FT DISULFID 47 111 BY SIMILARITY.  
 FT DISULFID 149 205 N-LINKED (GLCNAc...) (POTENTIAL).  
 FT CARBOHYD 51 51 D->G (IN DQB1\*0202).  
 FT VARIANT 167 167 /FtId=VAR\_008115.  
 FT FT  
 FT CONFLICT 69 69 I->V (IN REF. 2 AND 3).  
 SQ SEQUENCE 261 AA; 29733 MM; 8957ACD3E93DD105 CRC64;  
 Query Match 55.6%; Score 5; DB 1; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2 STPVA 6  
 Db 26 STPVA 30  
 RESULT 25  
 HB24\_HUMAN STANDARD: PRT; 261 AA.  
 ID HB24\_HUMAN P01920;  
 AC P01920;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DQ(3) BETA CHAIN PRECURSOR  
 (CLONE II-102).  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84070816; PubMed=6316358;  
 RA Larhammar D., Hyldig-Nielsen J.J., Servenius B., Andersson G.,  
 Rask L., Peterson P.A.;  
 RT "Exon-intron organization and complete nucleotide sequence of a human  
 major histocompatibility antigen DC beta gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 80:7313-7317(1983).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87250501; PubMed=3036828;  
 RA Jonsson A.K., Hyldig-Nielsen J.J., Servenius B., Larhammar D.,  
 Andersson G., Joergensen F., Peterson P.A., Rask L.;  
 RT "Class II genes of the human major histocompatibility complex.  
 Comparisons of the DQ and DX alpha and beta genes.";  
 RL J. Biol. Chem. 262:8767-8777(1987).  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC CC  
 CC EMBL: K01499; AAA98746.1; -;  
 CC EMBL: M29613; -; NOT\_ANNOTATED\_CDS.  
 CC EMBL: M29616; -; NOT\_ANNOTATED\_CDS.  
 CC PIR: A02234; HLHUI.  
 CC PIR: C29312; C29312.  
 CC HSSP: P13760; 2SERB.  
 CC InterPro: IPR000353; -;  
 CC InterPro: IPR000495; -;  
 CC InterPro: IPR003006; -;  
 CC Pfam: PF00069; MHC\_II\_beta; 1.

PFam: PF00047; 19; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 KW MHC II; Transmembrane; Glycoprotein; Signal.  
 FT SIGNAL 1 32  
 FT CHAIN 33 261  
 FT DOMAIN 33 261  
 FT DOMAIN 127 220  
 FT DOMAIN 221 230  
 FT TRANSMEM 231 251  
 FT DOMAIN 252 261  
 FT DISULFID 47 111  
 FT DISULFID 149 205  
 FT CARBOHYD 51 51  
 SQ SEQUENCE 261 AA; 29870 MW; 802CFE6C08D04AF1 CRC64; (POTENTIAL).  
 Query Match 55.6%; Score 5; DB 1; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 STPVA 6  
 Db 26 STPVA 30  
 RESULT 26  
 GP3D\_CHLTR STANDARD; PRT; 263 AA.  
 ID AC P10557; P08783; P08784; Q46431;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE VIRUENCE PLASMID PROTEIN PGP3-D (PROTEINS P-6/P-7).  
 OS Chlamydia trachomatis.  
 OC Plasmid pLG440, plasmid pCHL1, and plasmid pCTT1.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 ON NCBI\_Taxid=813;  
 RX STRAIN-L2/434/BU; PLASMID-pLG440;  
 RX MEDLINE-89013895; PubMed-2845228;  
 RA Commanducci M., Ricci S., Ratti G.;  
 RT "The structure of a plasmid of Chlamydia trachomatis believed to be required for growth within mammalian cells."  
 RL Mol. Microbiol. 2:531-538(1988).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-L1/440/LN; PLASMID-pLG440;  
 RX MEDLINE-88233998; PubMed-2836808;  
 RA Halt C., Ward M.E., Clarke I.N.;  
 RT "Analysis of the entire nucleotide sequence of the cryptic plasmid of Chlamydia trachomatis serovar L1. Evidence for involvement in DNA replication."  
 RL Nucleic Acids Res. 16:4053-4067(1988).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-G0/86 / SEROTYPE D; PLASMID-pCHL1;  
 RX MEDLINE-90301796; PubMed-2194229;  
 RA Commanducci M., Ricci S., Cevenini R., Ratti G.;  
 RT "Diversity of the Chlamydia trachomatis common plasmid in biovars with different pathogenicity."  
 RL Plasmid 23:149-154(1990).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SEROTYPE B; PLASMID-pCTT1;  
 RX MEDLINE-88177106; PubMed-3444855;  
 RA Sriprakash K.S., Macavoy E.S.;  
 RT "Characterization and sequence of a plasmid from the trachoma biovar of Chlamydia trachomatis."  
 RL Plasmid 18:205-214(1987).  
 RN [5]  
 RP SEQUENCE OF 1-10.

RC STRAIN-L2/434/BU;  
 RA Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,  
 RA Commanducci M., Christianen G., Birkelund S., Vitreou E., Ratti G.,  
 RA Pallini V.;  
 RL Submitted (SEP-1994) to the SWISS-PROT data bank.  
 CC -1- MISCELLANEOUS: PGP3-D IS REQUIRED FOR GROWTH WITHIN MAMMALIAN CELLS.  
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF PLASMID pLG440.  
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 126.  
 CC -1- CAUTION: REF.4 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A FRAMESHIFT IN POSITION 38.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X07547; CAA30423.1; -;  
 DR EMBL: X06707; CAA29894.1; -;  
 DR EMBL: J03321; AAA91571.1; -;  
 DR EMBL: M19487; AAB02589.1; ALT\_FRAME.  
 DR PIR: S01922; S01922.  
 DR PIR: S00790; S00790.  
 DR PIR: S00791; S00791.  
 KW plasmid.  
 FT INITMET 0 0  
 FT VARIANT 11 11  
 FT VARIANT 60 60  
 FT VARIANT 85 85  
 FT VARIANT 89 89  
 FT VARIANT 108 108  
 FT VARIANT 137 137  
 FT VARIANT 188 188  
 FT VARIANT 190 190  
 FT VARIANT 200 200  
 FT VARIANT 209 209  
 FT VARIANT 211 211  
 SQ SEQUENCE 263 AA; 27791 MW; A3FF2469B1C89703 CRC64;  
 Query Match 55.6%; Score 5; DB 1; Length 263;  
 Best Local Similarity 100.0%; Pred. No. 90;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 STPVA 6  
 Db 39 STPVA 43  
 RESULT 27  
 PSBP\_FRIAG STANDARD; PRT; 264 AA.  
 ID AC Q49080;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE OXYGEN-EVOLVING ENHANCER PROTEIN 2, CHLOROPLAST PRECURSOR (OEE2) (23 KDA SUBUNIT OF OXYGEN EVOLVING SYSTEM OF PHOTOSYSTEM II) (OEC 23 KDA DE SUBUNIT) (23 KDA THYLAKOID MEMBRANE PROTEIN).  
 GN PSBP.  
 OS Eritillaria agrestis.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Liliales; Liliaceae; Eritillaria.  
 ON NCBI\_Taxid=64177;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Panico E., Baysdorfer C.;  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MAY BE INVOLVED IN THE REGULATION OF PHOTOSYSTEM II.

CC -1- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE; ASSOCIATED  
 CC WITH THE PHOTOSYSTEM II COMPLEX.  
 CC -1- SIMILARITY: BELONGS TO THE PSBO FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF037458; AAC04809.1; -  
 CC DR Mendel: 28047; Frlag:PsbP;28047.  
 CC DR InterPro: IPR002683; -  
 CC DR Pfam: PF01789; PsbP; 1.  
 CC DR Photosynthesis; Photosystem II; Chloroplast; Transit peptide;  
 CC Thylakoid membrane.  
 CC TRANSIT 1 78 CHLOROPLAST (BY SIMILARITY).  
 CC FT CHAIN 79 264 OXYGEN-EVOLVING ENHANCER PROTEIN 2.  
 CC SEQUENCE 264 AA: 28111 MW: 1E214D433C95809B CRC64.  
 CC -----  
 CC Query Match 55.6%; Score 5; DB 1; Length 264;  
 CC Best Local Similarity 100.0%; Pred. No. 90;  
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC Oy 1 VSTPV 5  
 CC Db 194 VSTPV 198  
 CC -----  
 CC RESULT 28  
 CC HB2X\_HUMAN STANDARD: PRT; 268 AA.  
 CC ID HB2X\_HUMAN  
 CC AC P05538;  
 CC DT 01-NOV-1988 (Rel. 09, Created)  
 CC DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 CC DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 CC DE HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DX BETA CHAIN PRECURSOR.  
 CC OS Homo sapiens (Human).  
 CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 CC OX NCBI\_Taxid=9606;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE-87250501; PubMed-3036828;  
 CC RA Jonsson A.K., Hyldig-Nielsen J.J., Serenius B., Larhammar D.,  
 CC Andersson G., Joergensen F., Peterson P.A., Rask L.;  
 CC RT "Class II genes of the human major histocompatibility complex.  
 CC RT Comparisons of the DQ and DX alpha and beta genes.";  
 CC RL J. Biol. Chem. 262:8767-8777(1987).  
 CC RN [2]  
 CC RP SEQUENCE OF 38-125 FROM N.A.  
 CC RX MEDLINE-85216510; PubMed-3858830;  
 CC RA Okada K., Bose J.M., Prentice H., Spies T., Mengler R., Auffray C.,  
 CC Lillie J.W., Grossberger D., Strominger J.L.;  
 CC RT "Gene organization of DC and DX subregions of the human major  
 CC RT histocompatibility complex.";  
 CC RL Proc. Natl. Acad. Sci. U.S.A. 82:3410-3414(1985).  
 CC CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M29614; -; NOT\_ANNOTATED\_CDS.  
 CC DR EMBL: M29615; -; NOT\_ANNOTATED\_CDS.  
 CC DR EMBL: M11136; -; NOT\_ANNOTATED\_CDS.  
 CC DR PIR: D29312; D29312.

DR HSSP: P13760; 2SEB.  
 DR InterPro: IPR000353; -  
 DR InterPro: IPR000495; -  
 DR InterPro: IPR003006; -  
 DR Pfam: PF00969; MHC\_II\_beta; 1.  
 DR Pfam: PF00047; Ig; 1.  
 DR PROSITE: PS00290; IG-MHC; 1.  
 CC MHC II; Transmembrane; Glycoprotein; Signal.  
 CC FT SIGNAL 1 32  
 CC FT CHAIN 33 268  
 CC FT DOMAIN 33 126  
 CC FT DOMAIN 127 229  
 CC FT TRANSMEM 230 250  
 CC FT DOMAIN 251 268  
 CC FT DISULFID 47 110  
 CC FT DISULFID 148 204  
 CC FT CARBOHYD 51 51  
 CC SEQUENCE 268 AA: 30386 MW: 2746ED6CC5D44AF2 CRC64.  
 CC -----  
 CC Query Match 55.6%; Score 5; DB 1; Length 268;  
 CC Best Local Similarity 100.0%; Pred. No. 91;  
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC Oy 2 STPVA 6  
 CC Db 26 STPVA 30  
 CC -----  
 CC RESULT 29  
 CC YSM4\_CAEEL STANDARD: PRT; 271 AA.  
 CC ID YSM4\_CAEEL  
 CC AC Q10124;  
 CC DT 01-FEB-1996 (Rel. 33, Created)  
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 CC DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 CC DE HYPOTHETICAL 30.6 KDA PROTEIN F52C9.4 IN CHROMOSOME III.  
 CC GN F52C9.4.  
 CC OS Caenorhabditis elegans.  
 CC CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 CC Rhabditidae; Peloderinae; Caenorhabditis.  
 CC OX NCBI\_Taxid=6239;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN-BRISTOL N2;  
 CC RA Faveillo T.;  
 CC RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.  
 CC CC -1- SIMILARITY: TO CHIMPANZEE PROTEIN GOR, YEAST YGR276C AND X.LAEVIS  
 CC XPMC2.  
 CC CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U39850; AAA81059.1; -  
 CC DR WormPep: F52C9.4; CE01962.  
 CC DR InterPro: IPR000520; -  
 CC DR Pfam: PF00929; Exonuclease; 1.  
 CC KM Hypothetical protein.  
 CC SQ SEQUENCE 271 AA: 30613 MW: 4BA6FC40603D583 CRC64.  
 CC -----  
 CC Query Match 55.6%; Score 5; DB 1; Length 271;  
 CC Best Local Similarity 100.0%; Pred. No. 92;  
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC Oy 1 VSTPV 5  
 CC Db 11111

Db 87 VSTPV 91

RESULT 30  
Y4OR\_RHISN  
ID Y4OR\_RHISN STANDARD; PRT; 277 AA.  
AC p55603;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DE PROBABLE ABC TRANSPORTER PERMEASE PROTEIN Y4OR.  
GN Y4OR.  
OS Rhizobium sp. (strain NGR34).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97305956; PubMed=9163424;  
RA Freilberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A., Perret X.;  
RT "Molecular basis of symbiosis between Rhizobium and legumes."  
RL Nucle 387:394-401(1997).  
CC -1- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT  
CC SYSTEM Y4OPQRS. THIS SYSTEM PROBABLY TRANSPORTS A SUGAR-LIKE  
CC MOLECULE. PROBABLY RESPONSIBLE FOR THE TRANSLLOCATION OF THE  
CC SUBSTRATE ACROSS THE MEMBRANE.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
CC (POTENTIAL).  
CC -1- SIMILARITY: WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-  
CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE MALFG  
CC SUBFAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AEO00089; AAB91804.1; -  
DR InterPro: IPR000515; -  
DR Pfam: PF00528; BPD\_transp. 1.  
DR PROSITE: PS00402; BPD\_TRANSPP\_INN\_MEMBR. 1.  
KM Hypothetical protein; Inner membrane; Transmembrane; Transport;  
KM plasmid.  
FT TRANSMEM 15 35 POTENTIAL.  
FT TRANSMEM 79 99 POTENTIAL.  
FT TRANSMEM 110 130 POTENTIAL.  
FT TRANSMEM 140 160 POTENTIAL.  
FT TRANSMEM 189 209 POTENTIAL.  
FT TRANSMEM 214 234 POTENTIAL.  
FT TRANSMEM 242 262 POTENTIAL.  
SQ SEQUENCE 277 AA; 30576 MW; 48B863DD615B1D3D CRC64;

Query Match 55.6%; Score 5; DB 1; Length 277;  
Best Local Similarity 100.0%; Pred. 94;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
Db 252 VSTPV 256

RESULT 31  
PPV\_DROME  
ID PPV\_DROME STANDARD; PRT; 303 AA.  
AC Q27884; Q9M328;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE SERINE/THREONINE PROTEIN PHOSPHATASE pp-V (EC 3.1.3.16).  
GN PPV OR PPV6A OR CG12217.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.. AND CHARACTERIZATION.  
RC STRAIN=OREGON-R; TISSUE=Head, and Eye Imaginal disk;  
RX MEDLINE=94038668; PubMed=8223492;  
RA Mann D.J., Dombrodt V., Cohen P.T.W.;  
RT "Drosophila protein phosphatase V functionally complements a SIT4  
RT mutant in Saccharomyces cerevisiae and its amino-terminal region can  
RT confer this complementation to a heterologous phosphatase catalytic  
RT domain."  
RL EMO J. 12:4833-4842(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA April J.F., Agbayani A., An H.-J., Andrews-Plankoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borikova D., Botchan M.R., Bouck J.P., Broksstein P., Brothier P.,  
RA Burtils K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasner K.,  
RA Godler A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostli D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Gibbs R.A., Myers E.W., Rubin W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000)  
CC -1- FUNCTION: MAY BE INVOLVED IN CONTROLLING CELLULARIZATION OR IN  
CC REGULATING TRANSCRIPTION OF THE GENES INVOLVED IN THIS PROCESS.  
CC -1- CATALYTIC ACTIVITY: A PHOSPHOPROTEIN + H(2)O =  
CC A PROTEIN + ORTHOPHOSPHATE (THIS ENZYME IS SERINE/THREONINE  
CC SPECIFIC).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT HIGHEST LEVELS IN 2-4 H  
CC EMBRYOS.  
CC -1- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES. PP-V  
CC SUBFAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X75980; CAA53588.1; -  
 DR EMBL: AE003437; AAF46163.1; -  
 DR HSSP: P08129; IFCM. -  
 DR P1yase; FBgn003139; PPV.  
 DR InterPro: IPR000934; -  
 DR Pfam: PF00149; STphosphatase; 1.  
 DR PRINTS: PR00114; STPHPHASE.  
 DR PROSITE: PS00125; SER\_THR\_PHOSPHATASE; 1.  
 KW Hydrolase; Iron; Manganese.  
 FT METAL 51 51 IRON (BY SIMILARITY).  
 FT METAL 53 53 IRON (BY SIMILARITY).  
 FT METAL 79 79 IRON AND MANGANESE (BY SIMILARITY).  
 FT METAL 111 111 MANGANESE (BY SIMILARITY).  
 FT ACT SITE 112 112 GENERAL ACID (BY SIMILARITY).  
 FT METAL 161 161 MANGANESE (BY SIMILARITY).  
 FT METAL 235 235 MANGANESE (BY SIMILARITY).  
 SQ SEQUENCE 303 AA: 34759 MW; 5269f971416e55ec CRC64;

Query Match 55.6%; Score 5; DB 1; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 5; Conservative 0;

OY 1 VSTPV 5  
 |||||  
 DB 42 VSTPV 46

RESULT 32  
 PPP6\_HUMAN STANDARD: PRT; 305 AA.  
 ID AC 000743;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE SERINE/THREONINE PROTEIN PHOSPHATASE 6 (EC 3.1.3.16) (PP6).  
 GN PPP6C OR PPP6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97165573; PubMed=9013334;  
 RA Bastians H., Ponstingl H.;  
 RT The novel human protein serine/threonine phosphatase 6 is a  
 RT functional homologue of budding yeast Slf4p and fission yeast pel,  
 RT which are involved in cell cycle regulation.";  
 RL J. Cell Sci. 109:2865-2874(1996).  
 CC -1- FUNCTION: MAY FUNCTION IN CELL CYCLE REGULATION.  
 CC -1- CATALYTIC ACTIVITY: A PHOSPHOPROTEIN + H(2)O =  
 CC A PROTEIN + ORTHOPHOSPHATE (THIS ENZYME IS SERINE/THREONINE  
 CC SPECIFIC).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION LEVELS IN TESTIS, HEART,  
 CC AND SKELETAL MUSCLE AND LOWEST IN PLACENTA, LUNG, AND KIDNEY.  
 CC -1- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES. PP-V  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X92972; CAA63549.1; -  
 DR HSSP: P08129; IFCM.  
 DR MIM: 300141; -  
 DR InterPro: IPR000934; -  
 DR Pfam: PF00149; STphosphatase; 1.  
 DR PRINTS: PR00114; STPHPHASE.  
 DR PROSITE: PS00125; SER\_THR\_PHOSPHATASE; 1.  
 KW Hydrolase; Iron; Manganese.  
 FT METAL 53 53 IRON (BY SIMILARITY).  
 FT METAL 55 55 IRON (BY SIMILARITY).  
 FT METAL 81 81 IRON AND MANGANESE (BY SIMILARITY).  
 FT METAL 113 113 MANGANESE (BY SIMILARITY).  
 FT ACT SITE 114 114 GENERAL ACID (BY SIMILARITY).  
 FT METAL 163 163 MANGANESE (BY SIMILARITY).  
 FT METAL 237 237 MANGANESE (BY SIMILARITY).  
 SQ SEQUENCE 305 AA: 35144 MW; 35E95DB6C9A7CF4 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 5; Conservative 0;

OY 1 VSTPV 5  
 |||||  
 DB 44 VSTPV 48

RESULT 33  
 PPP6\_RAT STANDARD: PRT; 305 AA.  
 ID AC 064620;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE SERINE/THREONINE PROTEIN PHOSPHATASE 6 (EC 3.1.3.16) (PP6) (PROTEIN  
 DE PHOSPHATASE V) (PP-V).  
 GN PPP6C OR PPV.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Testis, and Brain;  
 RX MEDLINE=94357899; PubMed=8077208;  
 RA Becker W., Kentrup H., Klump S., Schultz J.E., Joost H.G.;  
 RT Molecular cloning of a protein serine/threonine phosphatase  
 RT containing a putative regulatory tetrapeptide repeat domain.";  
 RL J. Biol. Chem. 269:22586-22592(1994).  
 CC -1- FUNCTION: MAY FUNCTION IN CELL CYCLE REGULATION.  
 CC -1- CATALYTIC ACTIVITY: A PHOSPHOPROTEIN + H(2)O =  
 CC A PROTEIN + ORTHOPHOSPHATE (THIS ENZYME IS SERINE/THREONINE  
 CC SPECIFIC).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS FOUND IN SPLEEN, BRAIN AND  
 CC LUNG.  
 CC -1- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES. PP-V  
 CC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

KW Hydroxylase; Iron; Manganese.  
 FT METAL 53 53 IRON (BY SIMILARITY).  
 FT METAL 55 55 IRON (BY SIMILARITY).  
 FT METAL 81 81 IRON AND MANGANESE (BY SIMILARITY).  
 FT METAL 113 113 MANGANESE (BY SIMILARITY).  
 FT ACT\_SITE 114 114 GENERAL ACID (BY SIMILARITY).  
 FT METAL 163 163 MANGANESE (BY SIMILARITY).  
 FT METAL 237 237 MANGANESE (BY SIMILARITY).  
 SQ SEQUENCE 305 AA; 35106 MW; A91F993FBD8F110 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 305;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 |||||  
 DB 44 VSTPV 48

RESULT 34  
 POLG\_HCVH7 STANDARD; PRT; 309 AA.  
 ID POLG\_HCVH7  
 AC P27955;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE GENOME POLYPROTEIN [CONTAINS: MATRIX PROTEIN (ENVELOPE PROTEIN M);  
 MAJOR ENVELOPE PROTEIN E; NONSTRUCTURAL PROTEIN NS1 (FRAGMENT).  
 OS Hepatitis C virus (isolate HCT27) (HCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage: Flaviviridae;  
 OC Hepacivirus.  
 NC NCB1\_Taxid=11109;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91112009; PubMed=1846505;  
 RA Weiner A.J., Brauer M.J., Rosenblatt J., Richman K.H., Tung J.,  
 RA Crawford K., Bonino F., Saracco G., Choo Q.-L., Houghton M.,  
 RA Han J.H.;  
 RT "Variable and hypervariable domains are found in the regions of HCV  
 RT corresponding to the flavivirus envelope and NS1 proteins and the  
 RT pestivirus envelope glycoproteins.";  
 RL Virology 180:842-848(1991).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X51333; CAA37293.1; -  
 DR InterPro: IPR002519; -  
 DR InterPro: IPR002521; -  
 DR InterPro: IPR002531; -  
 DR Pfam: PF01560; HCV\_NS1; 1.  
 DR Pfam: PF01542; HCV\_core; 1.  
 DR Pfam: PF01539; HCV\_env; 1.  
 KW Polypeptin. Glycoprotein. Coat protein. Envelope protein;  
 KW Transmembrane; Nonstructural protein.  
 FT NON\_TER 1 1  
 FT CHAIN <1 63 MATRIX PROTEIN (POTENTIAL).  
 FT CHAIN 64 255 MAJOR ENVELOPE PROTEIN E (POTENTIAL).  
 FT CHAIN 256 >309 NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).  
 FT CARBOHYD 68 68  
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT NON\_TER 309 309

SQ SEQUENCE 309 AA; 32922 MW; 6E58E9C3D0B9EA9 CRC64;  
 Query Match 55.6%; Score 5; DB 1; Length 309;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
 |||||  
 DB 113 PVAPT 117

RESULT 35  
 ACCO\_MALDO STANDARD; PRT; 314 AA.  
 ID ACCO\_MALDO  
 AC 000985; 024063;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 1-AMINOCYCLOPROPANE-1-CARBOXYLATE OXIDASE (EC 1.-.-.-) (ACC OXIDASE)  
 OS ETHYLENE-FORMING ENZYME (EFE) (PROTEIN AP4 OR PAE12).  
 DE Malus domestica (Apple) (Malus sylvestris).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I;  
 OC Rosales; Rosaceae; Malus.  
 NC NCB1\_Taxid=3750;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. GOLDEN DELICIOUS; TISSUE=fruit;  
 RX MEDLINE=92322950; PubMed=1377961;  
 RA Ross G.S., Knighton M.L., Lay-Yee M.;  
 RT "An ethylene-related cDNA from ripening apples.";  
 RL Plant Mol. Biol. 19:231-238(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Dong J.G., Olsen D.B., Silverstone A., Yang S.F.;  
 RT "Sequence of a cDNA coding for a 1-aminocyclopropane-1-carboxylate  
 RT oxidase homolog from apple fruit.";  
 RL Plant Physiol. 98:1530-1531(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. GOLDEN DELICIOUS;  
 RA Castiglione S., Malerba M., Pirola B., Bianchetti R., Sala F.,  
 RA Ventura M., Pancaldi M., Sansavini S.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP CHARACTERIZATION, AND SEQUENCE OF 115-134.  
 RX MEDLINE=93028348; PubMed=1409700;  
 RA Dong J.G., Fernandez-Maculet J.C., Yang S.F.;  
 RT "Purification and characterization of  
 RT 1-aminocyclopropane-1-carboxylate oxidase from apple fruit.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:9789-9793(1992).  
 RN [5]  
 RP CHARACTERIZATION, AND PARTIAL SEQUENCE.  
 RX MEDLINE=93250487; PubMed=7763615;  
 RA Dupille E., Rombaldi C., Lelievre J.M., Cleyet-Marel J.C., Pech J.C.,  
 RA Lathe A.;  
 RT "Purification, properties and partial amino-acid sequence of 1-  
 RT aminocyclopropane-1-carboxylic acid oxidase from apple fruits.";  
 RL Planta 190:65-70(1993).  
 CC -1- CATALYTIC ACTIVITY: 1-AMINOCYCLOPROPANE-1-CARBOXYLATE + O2 -  
 CC ETHYLENE + HCN + CO(2) + 2 H(2)O.  
 CC -1- COFACTOR: IRON AND ASCORBATE.  
 CC -1- PATHWAY: LAST STEP IN THE BIOSYNTHESIS OF ETHYLENE.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING FRUIT RIPENING.  
 CC -1- INDUCTION: BY ETHYLENE AND BY WOUNDING.  
 CC -1- SIMILARITY: BELONGS TO THE IRON/ASCORBATE-DEPENDENT FAMILY OF  
 CC OXIDOREDUCTASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).

DR EMBL: X61390; CAA43662.1; -  
 DR EMBL: M81794; AAA33412.1; -  
 DR EMBL: Y14005; CAA74328.1; -  
 DR PIR: S22513; S22513.  
 DR Interpro: IPR002419; -  
 DR Pfam: PF00671; Fe\_Asc\_Oxidored; 1.  
 KW Fruit ripening; Ethylene biosynthesis; Oxidoreductase; Iron;  
 KW Vitamin C; Multigene family.  
 SQ SEQUENCE 314 AA; 35410 MW; BEADA64C7AD10E1E CRC64;

Query Match 55.6%; Score 5; DB 1; Length 314;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
 |||||  
 Db 308 STPVA 312

RESULT 36  
 SPLR\_NPVOP STANDARD: PRT; 321 AA.  
 ID SPLR\_NPVOP 065328; 012842; 012553;  
 AC 065328; 012842; 012553;  
 DT 01-NOV-1997 (Rel. 35; Created)  
 DT 01-NOV-1997 (Rel. 35; Last sequence update)  
 DT 01-NOV-1997 (Rel. 35; Last annotation update)  
 DE SPHEROIDIN-LIKE PROTEIN PRECURSOR (SPINDOLIN-LIKE PROTEIN) (37 KDA  
 DE GLYCOPROTEIN).  
 GN SLP OR GP37.  
 OS Orygia pseudotsugata multicausid polyhedrosis virus (OpMPV).  
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
 OC Nucleopolyhedrovirus.  
 OX NCBI\_Taxid=10450;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93100831; PubMed=8380088;  
 RA Gross C.H., Molgarnot G.M., Russell R.L., Pearson M.N., Rohmann G.F.;  
 RT "A 37-kilodalton glycoprotein from a baculovirus of Orygia  
 RT pseudotsugata is localized to cytoplasmic inclusion bodies";  
 RT J. Virol. 67:469-475(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97271300; PubMed=9126251;  
 RA Ahrens C.H., Russell R.R., Funk C.J., Evans J., Harwood S.,  
 RA Rohmann G.F.;  
 RT "The sequence of the Orygia pseudotsugata multicausid polyhedrosis virus  
 RT polyhedrosis virus genome";  
 RT Virology 229:381-399(1997).  
 RL -1- FUNCTION: COMPONENT OF THE VIRUS OCCLUSION BODIES, WHICH ARE LARGE  
 RL PROTEINACOUS STRUCTURES (POLYHEDRA) THAT PROTECT THE VIRUS FROM  
 RL THE OUTSIDE ENVIRONMENT FOR EXTENDED PERIODS UNTIL THEY ARE  
 RL INGESTED BY INSECT LARVAE (BY SIMILARITY).  
 RL -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED (BY SIMILARITY).  
 RL -1- SUBCELLULAR LOCATION: ASSOCIATED WITH THE PERIPHERY OF OCCLUSION  
 RL BODIES AND WITH THE INTERNAL MEMBRANES OF INFECTED CELLS.  
 RL -1- DEVELOPMENTAL STAGE: VERY LATE PHASE OF INFECTION.  
 RL -1- SIMILARITY: TO GBPV AND HAEPV SPINDOLINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).

DR EMBL: D13306; BAA02566.1; -

DR EMBL: U75930; AAC59068.1; -  
 KW Viral occlusion body; Signal; Glycoprotein; Late protein.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 321 SPHEROIDIN-LIKE PROTEIN.  
 FT CARBOHYD 192 192 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 321 AA; 36136 MW; B935809F06B56CA9 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 321;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
 |||||  
 Db 23 STPVA 27

RESULT 37  
 PF27\_MOUSE STANDARD: PRT; 323 AA.  
 ID PF27\_MOUSE P52875;  
 AC P52875;  
 DT 01-OCT-1996 (Rel. 34; Created)  
 DT 01-OCT-1996 (Rel. 34; Last sequence update)  
 DT 01-OCT-2000 (Rel. 40; Last annotation update)  
 DE TRANSMEMBRANE PROTEIN PFT27 (TPA REGULATED LOCUS PROTEIN).  
 GN TPRL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Teratocarcinoma;  
 RX MEDLINE=89076283; PubMed=3202867;  
 RA Akagi J., Nomiyama H., Setoyama C., Shinada K., Akagi M.;  
 RT "Messenger RNA expressed in mouse teratocarcinoma stem cells and  
 RT down-regulated by a tumor-promoting phorbol ester codes for a novel  
 RT transmembrane protein";  
 RL Biochem. Biophys. Res. Commun. 157:548-557(1988).  
 RL -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 RL -1- DEVELOPMENTAL STAGE: EXPRESSED IN UNDIFFERENTIATED MOUSE F9  
 RL TERATOCARCINOMA STEM CELLS BUT DISAPPEARING RAPIDLY AFTER  
 RL TREATMENT WITH A TUMOR-PROMOTING PHORBOL ESTER.  
 RL -1- SIMILARITY: BELONGS TO THE UPF0016 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).

DR EMBL: M23568; AAA40456.1; -  
 DR MGD: MGI:894407; Tpral.  
 DR Interpro: IPR001727; -  
 DR Pfam: PF01169; UPF0016; 1.  
 DR PROSITE: PS01214; UPF0016; 1.  
 KW Transmembrane.  
 FT TRANSMEM 16 36 POTENTIAL.  
 FT TRANSMEM 90 110 POTENTIAL.  
 FT TRANSMEM 127 147 POTENTIAL.  
 FT TRANSMEM 152 172 POTENTIAL.  
 FT TRANSMEM 228 248 POTENTIAL.  
 FT TRANSMEM 267 287 POTENTIAL.  
 FT TRANSMEM 299 319 POTENTIAL.  
 SQ SEQUENCE 323 AA; 34683 MW; 4C1D1201ADEFI333 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



OY 3 TPVAP 7  
 ID 73 TPVAP 77

RESULT 38  
 HMGL\_HUMAN STANDARD; PRT; 325 AA.  
 AC P35914;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE HYDROXYMETHYLGUTARYL-COA LYASE, MITOCHONDRIAL PRECURSOR (EC 4.1.3.4)  
 DE (HMG-COA LYASE) (HL) (3-HYDROXY-3-METHYLGUTARATE-COA LYASE).  
 GN HMGL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=93179448; PubMed=8440722;  
 RA Mitchell G.A., Robert M.-F., Hruz P.W., Wang S., Fontaine G.,  
 RA Benhke C.E., Mende-Mueller L.M., Schappert K., Lee C., Gibson K.M.,  
 RA Mziorko H.M.;  
 RT "3-hydroxy-3-methylglutaryl coenzyme A lyase (HL). Cloning of human  
 RT and chicken liver HL cDNAs and characterization of a mutation causing  
 RT human HL deficiency.";  
 RL J. Biol. Chem. 268:4376-4381(1993).  
 RN [2]  
 RP SEQUENCE OF 21-325 FROM N.A.  
 RX MEDLINE=96207305; PubMed=8617516;  
 RA Wang S.P., Robert M.-F., Gibson K.M., Wanders R.J.A., Mitchell G.A.;  
 RT "3-Hydroxy-3-methylglutaryl CoA lyase (HL): mouse and human HL gene  
 RT (HMGL) cloning and detection of large gene deletions in two  
 RT unrelated HL-deficient patients.";  
 RL Genomics 33:99-104(1996).  
 RN [3]  
 RP VARIANT HMG ARG-233.  
 RX MEDLINE=96394619; PubMed=8798725;  
 RA Roberts J., Mitchell G.A., Mziorko H.M.;  
 RT "Modeling of mutational responsible for human 3-hydroxy-3-  
 RT methylglutaryl-CoA lyase deficiency implicates histidine-233 as an  
 RT active site residue.";  
 RL J. Biol. Chem. 271:24604-24609(1996).  
 RN [4]  
 RP VARIANTS HMG GLN-41; GLU-42; GLY-42 AND HIS-42.  
 RX MEDLINE=98130531; PubMed=9463337;  
 RA Mitchell G.A., Ozand P.T., Robert M.-F., Ashmarina L., Roberts J.,  
 RA Gibson K.M., Wanders R.J., Wang S., Chevalier I., Ploechl E.,  
 RA Mziorko H.;  
 RT "HMG CoA lyase deficiency: identification of five causal point  
 RT mutations in codons 41 and 42, including a frequent Saudi Arabian  
 RT mutation, R11Q.";  
 RL Am. J. Hum. Genet. 62:295-300(1998).  
 RN [5]  
 RP CATALYTIC ACTIVITY: (S)-3-HYDROXY-3-METHYLGUTARYL-COA -  
 CC ACETYL-COA + ACETOACETATE.  
 CC [6]  
 CC PATHWAY: FINAL STEP OF KETONEGENESIS AND LEUCINE CATABOLISM.  
 CC [7]  
 CC SUBUNIT: HOMODIMER.  
 CC [8]  
 CC SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.  
 CC [9]  
 CC TISSUE SPECIFICITY: FIBROBLASTS, LIVER AND LYMPHOBLASTS.  
 CC [10]  
 CC DISEASE: DEFECTS IN HMGL ARE THE CAUSE OF AN AUTOSOMAL RECESSIVE  
 CC DISEASE KNOWN AS HYDROXYMETHYLGUTARICACIDURIA (HMG) WHICH CAN  
 CC LEAD TO HYPOGLYCEMIA AND COMA.  
 CC [11]  
 CC SIMILARITY: BELONGS TO THE HMG-COA LYASE FAMILY.  
 CC [12]  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC -----  
 DR EMBL: U49719; AAB19099.1; JOINED.  
 DR EMBL: U49712; AAB19099.1; JOINED.  
 DR EMBL: U49713; AAB19099.1; JOINED.  
 DR EMBL: U49714; AAB19099.1; JOINED.  
 DR EMBL: U49715; AAB19099.1; JOINED.  
 DR EMBL: U49716; AAB19099.1; JOINED.  
 DR EMBL: U49717; AAB19099.1; JOINED.  
 DR EMBL: U49718; AAB19099.1; JOINED.  
 DR PIR: A45470; A45470.  
 DR MIM: 246450; .  
 DR InterPro: IPR000138; .  
 DR InterPro: IPR000891; .  
 DR Pfam: PF00682; HMG-Like; 1.  
 DR Prosite: PS01062; HMG-COA LYASE; 1.  
 KW Lyase; Mitochondrion; Transit peptide; Disease mutation.  
 FT TRANSIT 1 27 MITOCHONDRION (BY SIMILARITY).  
 FT CHAIN 28 325 HYDROXYMETHYLGUTARYL-COA LYASE.  
 FT ACT\_SITE 266 266 BY SIMILARITY.  
 FT VARIANT 41 41 R -> Q (IN HMG).  
 FT VARIANT 42 42 D -> E (IN HMG).  
 FT VARIANT 42 42 /FTID=VAR.003745.  
 FT VARIANT 42 42 D -> G (IN HMG).  
 FT VARIANT 42 42 /FTID=VAR.003746.  
 FT VARIANT 70 70 D -> H (IN HMG).  
 FT VARIANT 70 70 /FTID=VAR.003747.  
 FT VARIANT 70 70 V -> L (IN HMG).  
 FT VARIANT 233 233 /FTID=VAR.003748.  
 FT VARIANT 233 233 /FTID=VAR.003749.  
 FT CONFLICT 243 243 H -> R (IN HMG).  
 FT CONFLICT 243 243 T -> A (IN REF. 2).  
 SQ SEQUENCE 325 AA; 34390 MW; 4D2B3F9210A67331 CRC64;

Query Match 55.6%; Score 5; DB 1; Length 325;  
 Best local Similarity 100.0%; Pred. No. 1; le+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 ID 51 VSTPV 55

RESULT 39  
 MODD\_MYCBO STANDARD; PRT; 325 AA.  
 ID MODD\_MYCBO  
 AC O30620;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE MOLYBDATE UPTAKE SECRETED PROTEIN PRECURSOR (ALA-PRO RICH 45/47 KDA  
 DE SECRETED PROTEIN) (IMMUNOGENIC PROTEIN MP32) (ANTIGEN MP1-32) (45-KDA  
 DE GLYCOPROTEIN) (FIBRONECTIN ATTACHMENT PROTEIN) (FAP-B).  
 GN MODD OR APA.  
 OS Mycobacterium bovis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OC NCBI\_TaxID=1765;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=BCG;  
 RA Zhao W., Schorey J.S., Bong-Waste M., Brown E.J., Ratliff T.L.;  
 RT "Identification, sequence and characterization of the M. bovis BCG  
 RT fibronectin attachment protein.";  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 CC [2]  
 CC MODACD FOR MOLYBDENUM (BY SIMILARITY).  
 CC [3]  
 CC SUBCELLULAR LOCATION: SECRETED.  
 CC [4]  
 CC SIMILARITY: BELONGS TO THE MODD FAMILY.  
 CC [5]  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
 CC EMBL: AF013569: AAB71842.1: -  
 CC Molybdenum; Transport; Glycoprotein; Antigen; Signal.

DR SIGNAL 1 39 BY SIMILARITY.  
 FT CHAIN 40 325 MOLYBDATE UPTAKE SECRETED PROTEIN.  
 FT CARBOHYD 49 49 O-LINKED (MAN. . .) (BY SIMILARITY).  
 FT CARBOHYD 57 57 O-LINKED (MAN. . .) (BY SIMILARITY).  
 FT CARBOHYD 66 66 O-LINKED (MAN. . .) (BY SIMILARITY).  
 FT CARBOHYD 316 316 O-LINKED (MAN. . .) (BY SIMILARITY).  
 FT SEQUENCE 325 AA: 32666 MW: D3419CA5547D91E9 CRC64:

Query Match 55.6%; Score 5; DB 1; Length 325;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
 DB 66 TPVAP 70

## RESULT 40

MODD\_MYCTU STANDARD: PRT: 325 AA.

AC 050906; 008062; (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE MOLYBDATE UPTAKE SECRETED PROTEIN PRECURSOR (ALA-PRO RICH 45/47 KDA

DE SECRETED PROTEIN) (IMMUNOGENIC PROTEIN MPT32) (ANTIGEN MPT-32) (45-KDA

DE GLYCOPROTEIN) (FIBRONECTIN ATTACHMENT PROTEIN) (FAP-B).

GN MODD OR APA OR RY1860 OR MTCY359.13.

OS Mycobacterium tuberculosis.

OC Bacteria; Filumetes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1773;

RA SEQUENCE FROM N.A.

RC STRAIN-H37RV; MEDLINE=96009758; PubMed=7558311;

RA Laqueyrie A., Miltzer P., Romain F., Elgmeler K., Cole S.,

RA Marchel G.;

RT "Cloning, sequencing, and expression of the apa gene coding for the

RT Mycobacterium tuberculosis 45/47-kilodalton secreted antigen

RL complex.";

RL Infect. Immun. 63:4003-4010(1995).

RN SEQUENCE FROM N.A.

RN [3]

RN [4]

RP PARTIAL SEQUENCE, AND CHARACTERIZATION.

RX MEDLINE=91099989; PubMed=1898899;

RA Nagai S., Wiker H.G., Harboe M., Kinomoto M.;

RT "Isolation and partial characterization of major protein antigens in

RT the culture fluid of Mycobacterium tuberculosis.";

RL Infect. Immun. 59:372-382(1991).

RN [5]

RP PARTIAL SEQUENCE, AND GLYCOSYLATION.

RX MEDLINE=95347792; PubMed=7622204;

RA Dobos K.M., Swiderex K., Khoo K.-H., Brennan P.J., Belisle J.T.;

RT "Evidence for glycosylation sites on the 45-kilodalton glycoprotein of

RT Mycobacterium tuberculosis.";

RL Infect. Immun. 63:2846-2853(1995).

RN [6]

RP PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.

RX MEDLINE=96196153; PubMed=8626314;

RA Dobos K.M., Khoo K.-H., Swiderex K.M., Brennan P.J., Belisle J.T.;

RT "Definition of the full extent of glycosylation of the 45-kilodalton

RT glycoprotein of Mycobacterium tuberculosis.";

RL J. Bacteriol. 178:2498-2506(1996).

CC - FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM

CC MODABC FOR MOLYBDENUM.

CC - SUBCELLULAR LOCATION: SECRETED.

CC - SIMILARITY: BELONGS TO THE MODD FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL: X80268: CAA56555.1; -

DR EMBL: X99258; CAA67645.1; -

DR EMBL: 283859; CAB06127.1; -

DR TubercuList; RV1860; -

KW Molybdenum; Transport; Glycoprotein; Antigen; Signal.

FT SIGNAL 1 39 MOLYBDATE UPTAKE SECRETED PROTEIN.

FT CHAIN 40 325 O-LINKED (MAN. . .).

FT CARBOHYD 49 49 O-LINKED (MAN. . .).

FT CARBOHYD 57 57 O-LINKED (MAN. . .).

FT CARBOHYD 66 66 O-LINKED (MAN. . .).

FT CARBOHYD 316 316 O-LINKED (MAN. . .).

FT SEQUENCE 325 AA: 32720 MW: 595D0455A97BED CRC64:

Query Match 55.6%; Score 5; DB 1; Length 325;  
 Best Local Similarity 100.0%; Pred. No. 1,1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
 DB 66 TPVAP 70

## RESULT 41

MODD\_XENLA STANDARD: PRT: 336 AA.

AC P21574; MEDLINE=98295987; PubMed=9634230;

DT 01-MAY-1991 (Rel. 18, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DE CYTOPLASMIC RNA-BINDING PROTEIN P56 (Y BOX BINDING PROTEIN-2) (Y-BOX

DE TRANSCRIPTION FACTOR) (MRNP4).

GN FRG2.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI\_TaxID=8355;

RN SEQUENCE FROM N.A.

RN [1]

RX MEDLINE-91062413; PubMed-2247479;  
 RA Tafuri S.R., Wolffe A.P.;  
 RT "Xenopus Y-box transcription factors: molecular cloning, functional  
 analysis and developmental regulation";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9028-9032(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE-92107999; PubMed-1729676;  
 RA Murray M.T., Schiller D.L., Franke W.M.;  
 RT "Sequence analysis of cytoplasmic mRNA-binding proteins of Xenopus  
 RT oocytes identifies a family of RNA-binding proteins";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:11-15(1992).  
 RN [3]  
 RP PARTIAL SEQUENCE.  
 RC TISSUE-Ovary;  
 RX MEDLINE-9232467; PubMed-1629179;  
 RA Deschamps S., Viel A., Garrigos M., Denis H., Le Maire M.;  
 RT "MRNP4, a major mRNA-binding protein from Xenopus oocytes is  
 RT identical to transcription factor FPG Y2";  
 RL J. Biol. Chem. 267:13799-13802(1992).  
 RN [4]  
 RP PARTIAL SEQUENCE.  
 RC TISSUE-Ovary;  
 RX MEDLINE-91224309; PubMed-1902800;  
 RA Deschamps S., Viel A., Denis H., Le Maire M.;  
 RT "Purification of two thermostable components of messenger  
 RT ribonucleoprotein particles (MRNPs) from Xenopus laevis oocytes,  
 RT belonging to a novel class of RNA-binding proteins";  
 RL FEBS Lett. 282:110-114(1991).  
 CC -1- FUNCTION: BINDS TO CCAAT-CONTAINING Y BOX OF THE HSP70 GENES.  
 CC SEEMS TO BE A NEGATIVE REGULATORY FACTOR. ALSO BINDS TO MRNA.  
 CC -1- SUBUNIT: POSSIBLY FORMS A HETERODIMER WITH P54 IN THE 6S AND 15S  
 CC MRNA-BINDING PARTICLES.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC, EITHER FREE OR ASSOCIATED WITH  
 CC RIBONUCLEOPROTEIN PARTICLES.  
 CC -1- TISSUE SPECIFICITY: TESTIS AND IMMATURE OOCYTES.  
 CC -1- PTM: PHOSPHORYLATION ACTIVATES IN VITRO RNA-BINDING.  
 CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M59454; AAA49716.1; -;  
 DR EIR: B38274; B38274;  
 DR HSSP: P32081; INMG;  
 DR TRANSFAC: T00294; -;  
 DR InterPro: IPR002059; -;  
 DR Pfam: PF00313; CSD; 1;  
 DR PRINTS: PR00050; COLDSHOCK;  
 DR PROSITE: PS00352; COLD\_SHOCK; 1;  
 KW Transcription regulation; DNA-binding; Nuclear protein;  
 KW RNA-binding; Phosphorylation.  
 FT DOMAIN 44 108 CSD.  
 FT CONFLICT 254 254 A -> T (IN REF. 1).  
 SQ SEQUENCE 336 AA; 37202 MW; 4AD5038769C6B84D CRC64;

Query Match 55.6%; Score 5; DB 1; Length 336;  
 Best Local Similarity 100.0%; Pred. No. 1, le+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
 I I I I I  
 DB 207 PVAPT 211

RESULT 42  
 RL3\_METHH

ID RL3\_METHH STANDARD; PRT; 337 AA.  
 AC 026110;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE 50S RIBOSOMAL PROTEIN L3P.  
 GN RPL3P OR MTH2.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanothermobacter.  
 OX NCBI\_TaxID=145262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DELTA H;  
 RX MEDLINE-98037514; PubMed-9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakey D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-T., Rice P., Nolling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics";  
 RL J. Bacteriol. 179:7135-7155(1997).  
 CC -1- SIMILARITY: BELONGS TO THE L3P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE000795; AB84521.1; -;  
 DR InterPro: IPR000597; -;  
 DR Pfam: PF00297; Ribosomal\_L3; 1;  
 DR PROSITE: PS00474; RIBOSOMAL\_L3; 1;  
 KW Ribosomal protein.  
 SQ SEQUENCE 337 AA; 37252 MW; 9BB3B8E0FD744A0B CRC64;

Query Match 55.6%; Score 5; DB 1; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 1, le+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 I I I I I  
 DB 67 VSTPV 71

RESULT 43  
 YMA3\_MYCBO STANDARD; PRT; 354 AA.  
 ID YMA3\_MYCBO  
 AC 002279;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE HYPOTHETICAL 38.1 KDA PROTEIN IN MAS 3' REGION.  
 DE MYCOTHERETICAL BOVIS.  
 OS Mycobacterium bovis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1765;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BG;  
 RX MEDLINE-92406887; PubMed-1527058;  
 RA Mathur M., Kolatukudy P.E.;  
 RT "Molecular cloning and sequencing of the gene for mycocerosic acid  
 RT synthase, a novel fatty acid elongating multifunctional enzyme, from  
 RT Mycobacterium tuberculosis var. bovis Bacillus Calmette-Guerin";  
 RL J. Biol. Chem. 267:19388-19395(1992).

CC -1- SIMILARITY: SOME, TO M.TUBERCULOSIS RV2285.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL, M95808; AAA25370.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 354 AA; 37956 MW; 2C77C1259BD3686A CRC64;  
  
Query Match 55.6%; Score 5; DB 1; Length 354;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 4 PVAPT 8  
| | | | |  
DB 284 PVAPT 288  
  
RESULT 44  
HIS8\_ECOLI STANDARD: PRT; 356 AA.  
ID HIS8\_ECOLI  
AC P06986; STANDB: PRT; 356 AA.  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (EC 2.6.1.9) (IMIDAZOLE ACETOL-  
DE PHOSPHATE TRANSAMINASE).  
CN HISC.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE=89094829; PubMed=3062174;  
RA Carliomagn M.S., Chiarloti L., Alfano P., Nappo A.G., Bruni C.B.;  
RT "Structure and function of the Salmonella typhimurium and Escherichia  
RT coli K-12 histidine operons".  
RL J. Mol. Biol. 203:565-606(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE=86059230; PubMed=2999081;  
RA Grisolia V., Carliomagn M.S., Nappo A.G., Bruni C.B.;  
RT "Cloning, structure, and expression of the Escherichia coli K-12 hisc  
RT gene".  
RL J. Bacteriol. 164:1317-1323(1985).  
RN [3]  
RP REVISIONS, SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE=94260549; PubMed=8201624;  
RA Jovanovic G., Kostic T., Jankovic M., Savic D.J.;  
RT "Nucleotide sequence of the Escherichia coli K12 histidine operon  
RT revalued".  
RL J. Mol. Biol. 239:433-435(1994).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12".  
RL Science 277:1453-1474(1997).  
RN [5]

RP SEQUENCE FROM N.A.  
RC STRAIN-K12;  
RX MEDLINE=97251358; PubMed=9097040;  
RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,  
RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,  
RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
RA Nakado S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T.,  
RA Saito N., Sempel G., Seki Y., Sivasubram S., Tagami H.,  
RA Takeda J., Takemoto K., Mada C., Yamamoto Y., Horiuchi T.;  
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome  
RT corresponding to the 40.1-50.0 min region on the linkage map".  
RL DNA Res. 3:379-392(1996).  
CC -1- CATALYTIC ACTIVITY: L-HISTIDINOL-PHOSPHATE + 2-OXOGUTARATE -  
CC 3-(IMIDAZOL-4-YL)-2-OXOPROPYL PHOSPHATE + GLUTAMATE.  
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
CC -1- PATHWAY: EIGHTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
CC AMINOTRANSFERASES.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL, X03416; CAA27150.1; -  
DR EMBL, X13462; CAA31813.1; -  
DR EMBL, U02071; AAA19743.1; -  
DR EMBL, AE000293; AAC75082.1; -  
DR EMBL, D90840; CAB21779.1; -  
DR PIR, A30270; XNECHC.  
DR Ecocore; EG10446; hisc.  
DR InterPro; IPR001917; -  
DR Pfam; PF00222; aminotran\_2; 1.  
DR PROSITE; PS00599; AA.TRANSFER.CLASS\_2; 1.  
KW Histidine biosynthesis; Transferase; Aminotransferase;  
KW Pyridoxal phosphate.  
FT BINDING 214 214 PYRIDOXAL PHOSPHATE (PROBABLE).  
FT FT 130 130 L -> P (IN REF. 1 AND 2).  
FT CONFLICT 149 149 V -> A (IN REF. 1 AND 2).  
SQ SEQUENCE 356 AA; 39360 MW; 056CBB3CF894083F CRC64;  
  
Query Match 55.6%; Score 5; DB 1; Length 356;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 2 STPVA 6  
| | | | |  
DB 246 STPVA 250  
  
RESULT 45  
HIS8\_SALTY STANDARD: PRT; 359 AA.  
ID HIS8\_SALTY  
AC P10369;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last annotation update)  
DE HISTIDINOL-PHOSPHATE AMINOTRANSFERASE (EC 2.6.1.9) (IMIDAZOLE ACETOL-  
DE PHOSPHATE TRANSAMINASE).  
CN HISC.  
OS Salmonella typhimurium.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Salmonella.  
OX NCBI\_TaxID=602;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-LT2;  
RX MEDLINE=89094829; PubMed=3062174;

RA Carlomagno M.S., Chiarlotti L., Alfano P., Nappo A.G., Bruni C.B.;  
 RT "Structure and function of the Salmonella typhimurium and Escherichia  
 coli K-12 histidine operons";  
 RL J. Mol. Biol. 203:585-606(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LT2;  
 RA Barnes W.M., Hussion R.N., Whittier R.;  
 RL Submitted (AUG-1989) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: L-HISTIDINOL-PHOSPHATE + 2-OXOGUTARATE =  
 CC 3-(IMIDAZOL-4-YL)-2-OXOPROPYL PHOSPHATE + GLUTAMATE;  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.  
 CC -1- PATHWAY: EIGHTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SIMILARITY: BELONGS TO CLASS-II OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
 CC AMINOTRANSFERASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial/  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; X13464; CAA31824.1; -;  
 CC DR EMBL; J01804; AAA88616.1; -;  
 CC DR PIR; J50158; XNEBHC.  
 CC DR StyGene; SG10159; h1sc.  
 CC DR InterPro; IPR001917; -;  
 CC DR Pfam; PF00222; aminotran\_2; 1.  
 CC DR PROSITE; PS00599; AA\_TRANSFER\_CLASS\_2; 1.  
 CC KM Histidine biosynthesis; Transferase; Aminotransferase;  
 CC PYridoxal phosphate.  
 CC FT BINDING 217 217 PYRIDOXAL PHOSPHATE (PROBABLE).  
 CC FT CONFLICT 148 164 DGRVVFVCSFNNPTGQ -> TAAKWCFAVPIILPQN  
 CC (IN REF. 2).  
 CC FT CONFLICT 226 226 C -> R (IN REF. 1).  
 CC FT CONFLICT 260 260 L -> S (IN REF. 1).  
 CC FT CONFLICT 264 265 GI -> ES (IN REF. 1).  
 CC FT CONFLICT 349 349 R -> P (IN REF. 1).  
 CC SQ SEQUENCE 359 AA; 39715 MW; 6241506A59CB2A7 CRC64;  
 CC -----  
 CC Query Match 55.6%; Score 5; DB 1; Length 359;  
 CC Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC OY 2 STPVA 6  
 CC Db 249 STPVA 253  
 CC -----  
 CC RESULT 46  
 CC SRK2-SPOLA STANDARD; PRT; 362 AA.  
 CC AC P42688;  
 CC DT 01-NOV-1995 (Rel. 32, Created)  
 CC DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 CC DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 CC DE TYROSINE-PROTEIN KINASE SRK2 (EC 2.7.1.112) (FRAGMENT).  
 CC GN SRK2.  
 CC OS Spongilla lacustris (Freshwater sponge).  
 CC OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;  
 CC OC Haplosclerida; Spongillidae; Spongilla.  
 CC OX NCBI\_TaxID=6055;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=92334872; PubMed=1378585;  
 CC RA Ottile S., Raulf F., Barnekow A., Hannig G., Scharf M.;  
 CC RT "Multiple src-related kinase genes, srk1-4, in the fresh water sponge  
 CC Spongilla lacustris";  
 CC RT Oncogene 7:1625-1630(1992).  
 CC -----

CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).  
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 CC DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial/  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; X61602; CAA43799.1; -;  
 CC DR HSSP; P00523; 2PTRK.  
 CC DR InterPro; IPR000719; -;  
 CC DR InterPro; IPR000980; -;  
 CC DR InterPro; IPR001245; -;  
 CC DR Pfam; PF00017; SH2; 1.  
 CC DR Pfam; PF00069; kinase; 1.  
 CC DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC DR PROSITE; PS50001; SH2; 1.  
 CC KM Transferase; Tyrosine-protein kinase; ATP-binding; SH2 domain;  
 CC KW phosphorylation.  
 CC FT NON\_TER 1  
 CC FT DOMAIN 1 70 SH2.  
 CC FT NP\_BIND 95 348 PROTEIN KINASE.  
 CC FT NP\_BIND 101 109 ATP (BY SIMILARITY).  
 CC FT BINDING 123 123 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 214 214 BY SIMILARITY.  
 CC SQ SEQUENCE 362 AA; 40937 MW; D8C549E904EBBD0 CRC64;  
 CC -----  
 CC Query Match 55.6%; Score 5; DB 1; Length 362;  
 CC Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC OY 2 STPVA 6  
 CC Db 117 STPVA 121  
 CC -----  
 CC RESULT 47  
 CC RROV\_REOVD STANDARD; PRT; 366 AA.  
 CC AC P03526;  
 CC DT 21-JUL-1986 (Rel. 01, Created)  
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 CC DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (SIGMA NS PROTEIN) (CORE  
 CC PROTEIN) (RNA-BINDING PROTEIN).  
 CC GN S3.  
 CC OS Reovirus (type 3 / strain Deering).  
 CC OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.  
 CC OX NCBI\_TaxID=10886;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=84015379; PubMed=6312421;  
 CC RA Richardson M.A., Furulich Y.;  
 CC RT "Nucleotide sequence of reovirus genome segment S3, encoding non-  
 CC structural protein sigma NS";  
 CC RT Nucleic Acids Res. 11:6399-6408(1983).  
 CC RN [2]  
 CC RP SEQUENCE OF 1-17 FROM N.A.  
 CC RX MEDLINE=83017876; PubMed=6927854;  
 CC RA Antczak J.B., Chmelo R.A., Pickup D.J., Joklik W.K.;  
 CC RT "Sequence at both termini of the 10 genes of reovirus serotype 3  
 CC (strain Deering)";  
 CC RT Virology 121:307-319(1982).  
 CC -----

CC -1- FUNCTION: THIS PROTEIN HAS A POLY(C)-DEPENDENT POLY(G) POLYMERASE  
CC ACTIVITY, AND BINDS SINGLE-STRANDED, BUT NOT DOUBLE-STRANDED, RNA.  
CC -1- SIMILARITY: 86.3% SIMILARITY TO REOVIRUS SEROTYPE 2 SIGMA NS  
CC PROTEIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X01627; CAA25768.1; -  
DR EMBL: J02330; AAA47280.1; -  
DR PIR: A04126; MNXRSD.  
DR InterPro: IPR002507; -  
DR Pfam: PF01518; PolyG\_pol; 1.  
KW Core protein; RNA-binding; Transferase; RNA-directed RNA polymerase.  
SQ SEQUENCE 366 AA; 41056 MW; 201A6223C4290DD7D CRC64;  
  
Query Match 55.6%; Score 5; DB 1; Length 366;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 4 PVAPT 8  
Db 189 PVAPT 193  
  
RESULT 48  
RPO\_REOVL STANDARD: PRT: 366 AA.  
ID P12002.  
AC 01-OCT-1989 (Rel. 12, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (SIGMA NS PROTEIN) (CORE  
DE PROTEIN) (RNA-BINDING PROTEIN).  
GN S3.  
OS Reovirus (type 2 / strain D5/Jones).  
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.  
OX NCBI\_TaxID:10885;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-88072071; PubMed-3686825;  
RA Wiener J.R., Joklik W.R.;  
RT "Comparison of the reovirus serotype 1, 2, and 3 S3 genome segments  
RT encoding the nonstructural protein sigma NS.";  
RL Virology 161:332-339(1987).  
CC -1- FUNCTION: THIS PROTEIN HAS A POLY(C)-DEPENDENT POLY(G) POLYMERASE  
CC ACTIVITY, AND BINDS SINGLE-STRANDED, BUT NOT DOUBLE-STRANDED, RNA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M18390; AAA47282.1; -  
DR PIR: B27401; MNXRST.  
DR InterPro: IPR002507; -  
DR Pfam: PF01518; PolyG\_pol; 1.  
KW Core protein; RNA-binding; Transferase; RNA-directed RNA polymerase.  
SQ SEQUENCE 366 AA; 41334 MW; EF9202C3A6FC5085 CRC64;

Oy 4 PVAPT 8  
Db 189 PVAPT 193  
  
RESULT 49  
RPO\_REOVL STANDARD: PRT: 366 AA.  
ID P07940.  
AC 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (SIGMA NS PROTEIN) (CORE  
DE PROTEIN) (RNA-BINDING PROTEIN).  
GN S3.  
OS Reovirus (type 1 / strain Lang).  
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.  
OX NCBI\_TaxID:10884;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-87025802; PubMed-3767989;  
RA George C.X., Atwater J.A., Samuel C.E.;  
RT "Biosynthesis of reovirus-specified polypeptides. Molecular cDNA  
RT cloning and nucleotide sequence of the reovirus serotype 1 Lang  
RT strain s3 mRNA which encodes the nonstructural RNA-binding protein  
RT sigma NS.";  
RL Biochem. Biophys. Res. Commun. 139:845-851(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-88072071; PubMed-3686825;  
RA Wiener J.R., Joklik W.R.;  
RT "Comparison of the reovirus serotype 1, 2, and 3 S3 genome segments  
RT encoding the nonstructural protein sigma NS.";  
RL Virology 161:332-339(1987).  
CC -1- FUNCTION: THIS PROTEIN HAS A POLY(C)-DEPENDENT POLY(G) POLYMERASE  
CC ACTIVITY, AND BINDS SINGLE-STRANDED, BUT NOT DOUBLE-STRANDED, RNA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M14325; AAA47273.1; -  
DR EMBL: M18389; AAA47281.1; -  
DR PIR: A25068; MNXRST.  
DR InterPro: IPR002507; -  
DR Pfam: PF01518; PolyG\_pol; 1.  
KW Core protein; RNA-binding; Transferase; RNA-directed RNA polymerase.  
FT CONFLICT 250 V -> I (IN REF. 2).  
SQ SEQUENCE 366 AA; 41188 MW; B91726A9251CC9A CRC64;  
  
Query Match 55.6%; Score 5; DB 1; Length 366;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 4 PVAPT 8  
Db 189 PVAPT 193  
  
RESULT 50  
Y85\_YEAST STANDARD: PRT: 366 AA.  
ID Y85\_YEAST  
AC 004806.  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL 42.4 KDA PROTEIN IN F4A4-HOR7 INTERGENIC REGION.

```

GN YMR251W OR YMR9920.05.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RA Hunt S., Bowman S., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE YGR154C/YKR076W/YMR251W (YEAST) /
CC YOG (E.COLI) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z48639; CAA88578.1; -.
CC DR SGD; S0004863; YMR251W.
CC KW Hypothetical protein.
CC SQ SEQUENCE 366 AA; 42403 MW; C0BF9FD3AB5C15B CRC64;

```

```

Query Match          55.6%; Score 5; DB 1; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 STPYA 6
   |||||
Db 109 STPYA 113

```

Search completed: August 15, 2001, 12:42:54  
 Job time: 484 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:33:55 ; Search time 58.01 Seconds  
(without alignments)  
20.527 Million cell updates/sec

Title: US-09-372-036-26

Perfect score: 9

Sequence: 1 VSRPVAPQTQ 9

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 425026 seqs, 132305027 residues

Word size : 0

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL\_16:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mmc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_protent:\*
- 12: sp\_unclassified:\*
- 13: sp\_vertebrate:\*
- 14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	100.0	478	2	Q9RNU3 listeria mo
2	9	100.0	478	2	Q03493 listeria mo
3	7	77.8	530	13	Q9DEY2 Q9DEY2 colutunix co
4	6	66.7	102	2	Q9H1T7 Q9H1T7 pseudomonas
5	6	66.7	209	5	Q9YX73 Q9YX73 dirosophila
6	6	66.7	237	10	Q9FU39 Q9FU39 oryza sativ
7	6	66.7	288	2	Q84512 Q84512 chlamydia t
8	6	66.7	289	3	Q9KYN4 Q9KYN4 streptomyc
9	6	66.7	293	2	Q9RW05 Q9RW05 delinococcus
10	6	66.7	447	2	Q86673 Q86673 streptomyc
11	6	66.7	460	3	Q9XTQ9 Q9XTQ9 caenorhabd
12	6	66.7	637	3	Q08484 Q08484 saccharomyc
13	6	66.7	652	5	Q9VW35 Q9VW35 dirosophila
14	6	66.7	748	3	Q9P5T7 Q9P5T7 neurospora
15	6	66.7	748	10	Q22907 Q22907 arabidopsis
16	6	66.7	750	3	Q9HFA4 Q9HFA4 candida alb
17	6	66.7	755	3	Q9VB49 Q9VB49 dirosophila
18	6	66.7	792	5	Q9VAC1 Q9VAC1 ephydattia f
19	6	66.7	819	5	Q9VVA8 Q9VVA8 dirosophila

20	6	66.7	904	5	Q9VEX7 Q9VEX7 dirosophila
21	6	66.7	912	5	Q9NHCI Q9NHCI dirosophila
22	6	66.7	912	5	Q9NHBP Q9NHBP dirosophila
23	6	66.7	916	5	Q9NHCC Q9NHCC dirosophila
24	6	66.7	916	5	Q9NHCO Q9NHCO dirosophila
25	6	66.7	1052	5	Q9VSS0 Q9VSS0 dirosophila
26	6	66.7	1094	4	Q00203 Q00203 homo sapien
27	6	66.7	1105	11	Q9Z1T1 Q9Z1T1 mus musculu
28	6	66.7	1653	2	Q9LBC3 Q9LBC3 streptococc
29	6	66.7	2037	13	Q9PVZ2 Q9PVZ2 xenopus lae
30	6	66.7	49	5	Q9Y1V2 Q9Y1V2 hydra magni
31	5	55.6	62	2	Q47006 Q47006 escherichia
32	5	55.6	75	2	Q9RRH1 Q9RRH1 delinococcus
33	5	55.6	82	5	Q17728 Q17728 caenorhabd
34	5	55.6	88	5	P91335 P91335 caenorhabd
35	5	55.6	94	14	Q67673 Q67673 porcine ade
36	5	55.6	95	11	Q9QYD6 Q9QYD6 mus musculu
37	5	55.6	95	14	Q68302 Q68302 hepatitis c
38	5	55.6	102	14	Q93516 Q93516 glycydus ha
39	5	55.6	108	13	Q93516 Q93516 glycydus ha
40	5	55.6	114	6	Q97949 Q97949 saginus oe
41	5	55.6	114	6	Q97935 Q97935 saginus oe
42	5	55.6	115	2	Q56724 Q56724 veillonella
43	5	55.6	115	5	Q9W2V1 Q9W2V1 dirosophila
44	5	55.6	115	13	Q90222 Q90222 agkistrodon
45	5	55.6	117	2	Q9X0K8 Q9X0K8 thermotoga
46	5	55.6	117	4	Q9NZE0 Q9NZE0 homo sapien
47	5	55.6	117	5	Q9VWMO Q9VWMO dirosophila
48	5	55.6	117	13	Q90221 Q90221 agkistrodon
49	5	55.6	121	5	Q9N3V1 Q9N3V1 caenorhabd
50	5	55.6	121	10	Q23980 Q23980 hordem vul
51	5	55.6	122	5	Q9YX28 Q9YX28 dirosophila
52	5	55.6	123	1	O59455 O59455 pyrococcus
53	5	55.6	126	7	Q29965 Q29965 homo sapien
54	5	55.6	126	7	Q78219 Q78219 homo sapien
55	5	55.6	132	2	O04169 O04169 zea mays. (m
56	5	55.6	133	2	O87645 O87645 methylococc
57	5	55.6	137	5	Q9N3U4 Q9N3U4 caenorhabd
58	5	55.6	138	14	Q9EBV4 Q9EBV4 foot-and-mo
59	5	55.6	138	14	Q9EBV2 Q9EBV2 foot-and-mo
60	5	55.6	138	14	Q9E8T3 Q9E8T3 foot-and-mo
61	5	55.6	141	1	Q9VAL8 Q9VAL8 aeropyrum p
62	5	55.6	141	3	O05690 O05690 saccharomyc
63	5	55.6	143	2	O07266 O07266 mycobacteri
64	5	55.6	148	6	Q9GKP9 Q9GKP9 bos taurus
65	5	55.6	149	14	Q68732 Q68732 hepatitis c
66	5	55.6	156	5	Q9N8J8 Q9N8J8 trypanosoma
67	5	55.6	156	14	O69124 O69124 human herpe
68	5	55.6	157	2	O49991 O49991 mycobacteri
69	5	55.6	158	14	O69129 O69129 human herpe
70	5	55.6	161	2	Q9FBC1 Q9FBC1 streptococc
71	5	55.6	161	10	Q9SSS3 Q9SSS3 arabidopsis
72	5	55.6	162	2	Q9K8Q7 Q9K8Q7 bacillus ha
73	5	55.6	163	5	Q9N3V0 Q9N3V0 caenorhabd
74	5	55.6	165	2	O52819 O52819 rhizobium l
75	5	55.6	166	4	Q9WRT0 Q9WRT0 homo sapien
76	5	55.6	166	10	Q9M4T9 Q9M4T9 triticum ae
77	5	55.6	170	2	Q9KCO1 Q9KCO1 bacillus ha
78	5	55.6	170	2	O05551 O05551 listeria mo
79	5	55.6	175	5	Q9VR96 Q9VR96 dirosophila
80	5	55.6	175	10	Q9XHF4 Q9XHF4 sorghum bic
81	5	55.6	177	10	O43478 O43478 hordem vul
82	5	55.6	179	10	Q9Z0M4 Q9Z0M4 arabidopsis
83	5	55.6	180	10	Q9W9M0 Q9W9M0 arabidopsis
84	5	55.6	185	5	O9V916 O9V916 dirosophila
85	5	55.6	185	10	Q9MA15 Q9MA15 arabidopsis
86	5	55.6	187	5	Q9N344 Q9N344 caenorhabd
87	5	55.6	191	2	Q9XCK6 Q9XCK6 streptococc
88	5	55.6	192	14	O81390 O81390 hepatitis c
89	5	55.6	193	10	Q9LKP5 Q9LKP5 zea mays (m
90	5	55.6	195	2	O9L5W8 O9L5W8 streptococc
91	5	55.6	197	2	Q9RKU2 Q9RKU2 streptomyc
92	5	55.6	199	2	Q9K531 Q9K531 pseudomonas

93	5	55.6	200	14	056247	056247 tobacco rat	166	5	55.6	315	5	018121	018121 caenorhabd
94	5	55.6	201	3	P87261	P87261 saccharomyc	167	5	55.6	315	5	09XWA9	09XWA9 caenorhabd
95	5	55.6	206	2	Q91651	Q91651 pseudomonas	168	5	55.6	318	2	066259	066259 actinobacil
96	5	55.6	206	2	Q9EX31	Q9EX31 streptomyc	169	5	55.6	318	2	005380	005380 actinobacil
97	5	55.6	209	5	Q9V794	Q9V794 drosophila	170	5	55.6	319	10	09SC08	09SC08 arabadopsi
98	5	55.6	210	2	Q9XCJ9	Q9XCJ9 streptococc	171	5	55.6	323	11	09R292	09R292 mus muscul
99	5	55.6	210	2	Q9RNU2	Q9RNU2 streptococc	172	5	55.6	324	5	09KEV8	09KEV8 caenorhabd
100	5	55.6	211	2	Q9XU58	Q9XU58 caenorhabd	173	5	55.6	324	5	022161	022161 caenorhabd
101	5	55.6	212	2	Q9S4U5	Q9S4U5 streptococc	174	5	55.6	327	10	09LD07	09LD07 oryza sativ
102	5	55.6	212	11	Q9ERR0	Q9ERR0 mus muscul	175	5	55.6	327	14	069145	069145 human herpe
103	5	55.6	215	2	Q51491	Q51491 pseudomonas	176	5	55.6	329	5	09VY11	09VY11 drosophila
104	5	55.6	218	5	Q9V6P3	Q9V6P3 drosophila	177	5	55.6	331	2	069818	069818 streptomyc
105	5	55.6	220	11	Q9Z1E5	Q9Z1E5 rattus norv	178	5	55.6	332	5	023156	023156 caenorhabd
106	5	55.6	223	2	Q9XCL1	Q9XCL1 streptococc	179	5	55.6	334	4	015008	015008 homo sapien
107	5	55.6	223	2	Q9KMN9	Q9KMN9 vibrio chol	180	5	55.6	334	4	09H7H8	09H7H8 homo sapien
108	5	55.6	224	5	Q9GPI4	Q9GPI4 drosophila	181	5	55.6	335	2	09RX81	09RX81 delinococc
109	5	55.6	224	2	Q9REJ3	Q9REJ3 streptococc	182	5	55.6	335	5	09NF70	09NF70 caenorhabd
110	5	55.6	224	7	Q29967	Q29967 homo sapien	183	5	55.6	335	5	09NF69	09NF69 caenorhabd
111	5	55.6	228	2	Q9RUB6	Q9RUB6 delinococc	184	5	55.6	335	5	09N6L3	09N6L3 caenorhabd
112	5	55.6	228	2	Q9S4S6	Q9S4S6 shigella fl	185	5	55.6	339	2	Q9RXE8	Q9RXE8 delinococc
113	5	55.6	236	10	Q9LM00	Q9LM00 pinus taeda	186	5	55.6	342	2	Q9Z5F8	Q9Z5F8 mycobacteri
114	5	55.6	239	14	Q9IRJ1	Q9IRJ1 turkey herp	187	5	55.6	343	8	Q9Z238	Q9Z238 trichophyto
115	5	55.6	241	2	Q9X6C8	Q9X6C8 streptococc	188	5	55.6	347	5	Q9VRJ9	Q9VRJ9 drosophila
116	5	55.6	241	5	Q9GPI3	Q9GPI3 drosophila	189	5	55.6	351	5	Q9VST8	Q9VST8 drosophila
117	5	55.6	241	11	Q9Z135	Q9Z135 rattus norv	190	5	55.6	351	10	023142	023142 arabadopsi
118	5	55.6	241	14	Q84617	Q84617 parametium	191	5	55.6	353	5	Q9W4Y8	Q9W4Y8 drosophila
119	5	55.6	244	5	Q9XZX6	Q9XZX6 leishmania	192	5	55.6	354	5	Q9NHR6	Q9NHR6 dictyostell
120	5	55.6	245	2	Q84764	Q84764 chlamydia t	193	5	55.6	355	4	006265	006265 homo sapien
121	5	55.6	250	7	Q19711	Q19711 homo sapien	194	5	55.6	356	2	Q9S5C6	Q9S5C6 escherichia
122	5	55.6	250	7	Q19714	Q19714 homo sapien	195	5	55.6	358	2	Q69960	Q69960 streptomyc
123	5	55.6	251	2	Q9K8B5	Q9K8B5 bacillus ha	196	5	55.6	359	2	Q9WMN3	Q9WMN3 synechococ
124	5	55.6	252	10	Q9LOL8	Q9LOL8 arabadopsi	197	5	55.6	359	8	Q37749	Q37749 cephalospor
125	5	55.6	255	7	Q19506	Q19506 homo sapien	198	5	55.6	362	4	Q9UHS5	Q9UHS5 homo sapien
126	5	55.6	256	2	Q9XG11	Q9XG11 streptococc	199	5	55.6	362	10	082054	082054 saccharum o
127	5	55.6	256	4	Q9NUR3	Q9NUR3 homo sapien	200	5	55.6	366	14	086313	086313 reovirus sp
128	5	55.6	258	2	Q9K410	Q9K410 streptomyc	201	5	55.6	366	14	086314	086314 reovirus sp
129	5	55.6	260	5	Q17625	Q17625 caenorhabd	202	5	55.6	366	14	086315	086315 reovirus sp
130	5	55.6	260	7	P79551	P79551 homo sapien	203	5	55.6	366	14	086316	086316 reovirus sp
131	5	55.6	261	3	Q9P390	Q9P390 neurospora	204	5	55.6	366	14	086317	086317 reovirus sp
132	5	55.6	261	7	Q30096	Q30096 homo sapien	205	5	55.6	366	14	086318	086318 reovirus sp
133	5	55.6	261	7	Q30099	Q30099 homo sapien	206	5	55.6	366	14	086319	086319 reovirus sp
134	5	55.6	261	7	Q30061	Q30061 homo sapien	207	5	55.6	366	14	086320	086320 reovirus sp
135	5	55.6	261	7	Q31633	Q31633 homo sapien	208	5	55.6	366	14	086323	086323 reovirus sp
136	5	55.6	264	7	Q29826	Q29826 homo sapien	209	5	55.6	366	14	086307	086307 reovirus sp
137	5	55.6	264	7	Q29970	Q29970 homo sapien	210	5	55.6	366	14	086308	086308 reovirus sp
138	5	55.6	267	5	Q16704	Q16704 caenorhabd	211	5	55.6	366	14	086309	086309 reovirus sp
139	5	55.6	267	14	087048	087048 semliki for	212	5	55.6	366	14	086310	086310 reovirus sp
140	5	55.6	269	7	Q30155	Q30155 homo sapien	213	5	55.6	366	14	086311	086311 reovirus sp
141	5	55.6	273	1	Q9HR52	Q9HR52 halobacteri	214	5	55.6	366	14	086312	086312 reovirus sp
142	5	55.6	273	2	Q05936	Q05936 pseudomonas	215	5	55.6	366	14	Q9ENM1	Q9ENM1 reovirus (t
143	5	55.6	276	2	P73137	P73137 synechocyst	216	5	55.6	367	14	Q41165	Q41165 parametium
144	5	55.6	276	5	Q94215	Q94215 caenorhabd	217	5	55.6	368	2	Q9LA74	Q9LA74 acinetobact
145	5	55.6	282	2	Q9ZCX6	Q9ZCX6 rickettsia	218	5	55.6	368	2	Q9FOT0	Q9FOT0 pseudomonas
146	5	55.6	286	2	Q9R640	Q9R640 mycobacteri	219	5	55.6	368	10	082288	082288 arabadopsi
147	5	55.6	288	2	Q9JXR2	Q9JXR2 neisseria m	220	5	55.6	372	4	Q12863	Q12863 homo sapien
148	5	55.6	293	2	Q9ZU08	Q9ZU08 rhodococcus	221	5	55.6	372	14	Q9PYV0	Q9PYV0 xestia c-n1
149	5	55.6	293	2	Q9XAV7	Q9XAV7 streptomyc	222	5	55.6	375	10	Q39235	Q39235 arabadopsi
150	5	55.6	296	10	Q22812	Q22812 streptococc	223	5	55.6	375	10	Q49349	Q49349 arabadopsi
151	5	55.6	297	5	Q9VD14	Q9VD14 drosophila	224	5	55.6	376	5	Q9VR07	Q9VR07 drosophila
152	5	55.6	298	2	Q9XY66	Q9XY66 streptomyc	225	5	55.6	377	11	Q9Z223	Q9Z223 cricetus
153	5	55.6	298	5	Q9VU90	Q9VU90 drosophila	226	5	55.6	377	11	Q9QXJ0	Q9QXJ0 rattus norv
154	5	55.6	299	2	Q9KPI8	Q9KPI8 vibrio chol	227	5	55.6	378	2	Q55307	Q55307 streptococc
155	5	55.6	302	10	Q9SAH7	Q9SAH7 arabadopsi	228	5	55.6	380	1	Q9HRS8	Q9HRS8 halobacteri
156	5	55.6	303	4	Q9UIC9	Q9UIC9 homo sapien	229	5	55.6	382	2	P77144	P77144 escherichia
157	5	55.6	304	14	Q9J4M8	Q9J4M8 human immun	230	5	55.6	382	5	Q9U3P1	Q9U3P1 caenorhabd
158	5	55.6	306	10	Q9SDN3	Q9SDN3 prunus dulc	231	5	55.6	383	14	Q04397	Q04397 epstein-bar
159	5	55.6	307	2	Q9M5Q3	Q9M5Q3 petunia hyb	232	5	55.6	384	4	Q9S977	Q9S977 homo sapien
160	5	55.6	307	5	Q25410	Q25410 synechocyst	233	5	55.6	385	5	Q9GY42	Q9GY42 leishmania
161	5	55.6	309	5	Q22135	Q22135 caenorhabd	234	5	55.6	385	6	Q9G1L4	Q9G1L4 micropotamo
162	5	55.6	310	11	Q9VUY0	Q9VUY0 drosophila	235	5	55.6	387	2	Q9LAT2	Q9LAT2 riemeirella
163	5	55.6	310	11	Q9EC47	Q9EC47 mus muscul	236	5	55.6	387	4	Q9LMT0	Q9LMT0 riemeirella
164	5	55.6	311	10	Q22805	Q22805 arabadopsi	237	5	55.6	387	4	Q9H8F7	Q9H8F7 homo sapien
165	5	55.6	314	10	Q9XKZ2	Q9XKZ2 pyrus pyrif	238	5	55.6	388	5	Q9VUD1	Q9VUD1 drosophila



239	5	55.6	392	4	Q9NP15	Q9np15 homo sapien	312	5	55.6	475	2	Q9EU22	Q9eu22 escherichia
240	5	55.6	392	5	Q9VM77	Q9vm77 drosophila	313	5	55.6	475	10	Q9LNM7	Q9lnm7 arabisdopsis
241	5	55.6	393	4	Q9UPH5	Q9uph5 homo sapien	314	5	55.6	476	5	P91292	P91292 caenorhabdi
242	5	55.6	393	14	O41549	O41549 human immun	315	5	55.6	476	5	Q03996	Q03996 plasmodium
243	5	55.6	394	5	Q9XU00	Q9xu00 caenorhabdi	316	5	55.6	476	13	Q9Y119	Q9y119 agkistrodon
244	5	55.6	395	6	Q77673	Q77673 chaetophrac	317	5	55.6	477	2	Q83494	Q83494 treponema p
245	5	55.6	397	4	Q9UPH4	Q9uph4 homo sapien	318	5	55.6	478	5	Q25684	Q25684 plasmodium
246	5	55.6	398	4	Q9UGL3	Q9ugl3 homo sapien	319	5	55.6	478	13	Q98995	Q98995 vipera lebe
247	5	55.6	400	13	Q9IAY3	Q9iay3 agkistrodon	320	5	55.6	479	13	Q9PM70	Q9pm70 agkistrodon
248	5	55.6	400	13	Q9IAY1	Q9iay1 agkistrodon	321	5	55.6	479	13	Q9PW78	Q9pw78 agkistrodon
249	5	55.6	400	13	Q9IAY0	Q9iay0 agkistrodon	322	5	55.6	480	13	Q90220	Q90220 agkistrodon
250	5	55.6	402	2	P76102	P76102 escherichia	323	5	55.6	481	2	Q9L408	Q9l408 clostridium
251	5	55.6	404	2	Q86687	Q86687 streptomyce	324	5	55.6	481	3	Q91505	Q91505 trimeresuru
252	5	55.6	406	3	Q9UVD9	Q9uud9 alternaria	325	5	55.6	481	13	Q57413	Q57413 trimeresuru
253	5	55.6	407	13	Q92032	Q92032 agkistrodon	326	5	55.6	482	5	O44196	O44196 caenorhabdi
254	5	55.6	408	4	Q9UGL4	Q9ugl4 homo sapien	327	5	55.6	482	5	Q9XV38	Q9xv38 acanthamoeb
255	5	55.6	410	2	O05446	O05446 mycobacteri	328	5	55.6	482	13	Q03598	Q03598 plasmodium
256	5	55.6	410	2	Q9F4K8	Q9f4k8 rhizobium l	329	5	55.6	482	13	Q9PVK9	Q9pvk9 agkistrodon
257	5	55.6	410	10	Q9MA98	Q9ma98 arabisdopsis	330	5	55.6	483	13	Q9IAB0	Q9iab0 agkistrodon
258	5	55.6	411	2	Q9F4K7	Q9f4k7 rhizobium l	331	5	55.6	485	5	Q9N4N9	Q9n4n9 caenorhabdi
259	5	55.6	411	5	Q9Y943	Q9y943 drosophila	332	5	55.6	486	4	Q9UCX4	Q9ucx4 homo sapien
260	5	55.6	411	14	Q81814	Q81814 hepatitis c	333	5	55.6	487	13	Q92119	Q92119 trimeresuru
261	5	55.6	413	13	Q9PW35	Q9pw35 agkistrodon	334	5	55.6	488	10	Q9LNX1	Q9lnx1 arabisdopsis
262	5	55.6	414	13	Q90391	Q90391 crocalus at	335	5	55.6	488	14	Q9WT03	Q9wt03 human herpe
263	5	55.6	414	13	Q90392	Q90392 crocalus at	336	5	55.6	491	2	Q9F518	Q9f518 alteromonas
264	5	55.6	414	14	O81339	O81339 hepatitis c	337	5	55.6	496	5	Q17537	Q17537 caenorhabdi
265	5	55.6	414	14	P89958	P89958 hepatitis c	338	5	55.6	498	4	Q9H7A2	Q9h7a2 homo sapien
266	5	55.6	414	14	P89959	P89959 hepatitis c	339	5	55.6	499	11	O88714	O88714 mus musculu
267	5	55.6	415	14	O81554	O81554 hepatitis c	340	5	55.6	499	14	Q72860	Q72860 human immun
268	5	55.6	416	2	O31444	O31444 bacillus su	341	5	55.6	500	2	P73537	P73537 synechocyst
269	5	55.6	418	2	Q9K9F3	Q9k9f3 bacillus ha	342	5	55.6	503	11	Q9ROM4	Q9rom4 mus musculu
270	5	55.6	418	4	O75066	O75066 homo sapien	343	5	55.6	503	11	Q9ESZ1	Q9esz1 mus musculu
271	5	55.6	418	4	Q9WU48	Q9wu48 homo sapien	344	5	55.6	504	2	Q9ZBN5	Q9zbn5 streptomyce
272	5	55.6	420	10	Q9ZOC4	Q9zoc4 arabisdopsis	345	5	55.6	505	4	O14395	O14395 homo sapien
273	5	55.6	421	2	Q90Z40	Q9jz40 neisseria m	346	5	55.6	505	4	Q9UNR8	Q9unr8 homo sapien
274	5	55.6	421	2	Q90J34	Q9ju34 neisseria m	347	5	55.6	505	13	Q73795	Q73795 glyoxydus ha
275	5	55.6	422	2	P96208	P96208 mycobacteri	348	5	55.6	508	5	Q93227	Q93227 caenorhabdi
276	5	55.6	423	1	Q93642	Q93642 halobacteri	349	5	55.6	508	6	Q9CMC8	Q9cmc8 fells silve
277	5	55.6	423	1	Q9HQT5	Q9hqt5 halobacteri	350	5	55.6	508	6	Q9CMC7	Q9cmc7 bison bison
278	5	55.6	423	2	Q49339	Q49339 mycobacteri	351	5	55.6	509	6	Q9GMC7	Q9gmc7 human papil
279	5	55.6	423	2	Q9U1M3	Q9u1m3 drosophila	352	5	55.6	509	14	O56939	O56939 human papil
280	5	55.6	428	2	O911M0	O911m0 pseudomonas	353	5	55.6	511	2	O52569	O52569 amycolatops
281	5	55.6	428	4	O9P245	O9p245 homo sapien	354	5	55.6	511	2	O52569	O52569 amycolatops
282	5	55.6	428	10	Q9FM92	Q9fm92 arabisdopsis	355	5	55.6	512	5	O76725	O76725 caenorhabdi
283	5	55.6	429	2	Q9Z4B1	Q9z4b1 escherichia	356	5	55.6	513	13	Q9YG84	Q9yg84 aeropyrum p
284	5	55.6	429	2	Q9K2H0	Q9k2h0 salmonella	357	5	55.6	514	10	Q9JLL6	Q9jll6 chlamydomon
285	5	55.6	434	2	Q9X6H2	Q9x6h2 streptococc	358	5	55.6	516	14	O12670	O12670 colobus mon
286	5	55.6	435	5	Q9VG46	Q9vg46 human herpe	359	5	55.6	521	14	P88804	P88804 human papil
287	5	55.6	437	14	O69146	O69146 human herpe	360	5	55.6	530	2	Q9KJPD	Q9kjd0 arthropacte
288	5	55.6	444	1	Q9HQ15	Q9hq15 halobacteri	361	5	55.6	530	2	Q9I008	Q9i008 pseudomonas
289	5	55.6	444	5	Q9VN63	Q9vn63 drosophila	362	5	55.6	534	1	O26459	O26459 methanobact
290	5	55.6	445	2	O45794	O45794 bacterioides	363	5	55.6	534	2	Q9KJ53	Q9kjs3 streptomyce
291	5	55.6	445	2	Q9K199	Q9k199 bacterioides	364	5	55.6	535	14	Q9QM6	Q9qm6 chlitia viru
292	5	55.6	446	5	Q9N3V2	Q9n3v2 caenorhabdi	365	5	55.6	535	14	Q9IV39	Q9iv39 human calic
293	5	55.6	448	5	Q9GNP9	Q9gnf9 drosophila	366	5	55.6	540	4	O60230	O60230 homo sapien
294	5	55.6	448	13	Q92019	Q92019 xenopus lae	367	5	55.6	542	5	O23446	O23446 caenorhabdi
295	5	55.6	451	2	Q9K801	Q9k801 bacillus ha	368	5	55.6	543	10	Q92188	Q92188 arabisdopsis
296	5	55.6	452	5	O27696	O27696 musca domes	369	5	55.6	543	10	Q43302	Q43302 arabisdopsis
297	5	55.6	455	2	Q914C1	Q914c1 pseudomonas	370	5	55.6	544	10	Q9LY09	Q9ly09 arabisdopsis
298	5	55.6	459	14	Q9ZHC0	Q9znc0 streptococc	371	5	55.6	547	13	Q91349	Q91349 xenopus lae
299	5	55.6	460	5	Q9IR53	Q9irs3 human papil	372	5	55.6	548	2	Q9ZEY1	Q9zey1 listeria mo
300	5	55.6	460	5	O74664	O74664 drosophila	373	5	55.6	548	2	Q9EXF6	Q9exf6 listeria mo
301	5	55.6	462	10	Q9LPW4	Q9lpw4 arabisdopsis	374	5	55.6	548	2	Q9XKV9	Q9xkv9 bombyx moti
302	5	55.6	464	2	P74828	P74828 spingomonas	375	5	55.6	548	2	Q9XKV9	Q9xkv9 bombyx moti
303	5	55.6	465	14	Q90729	Q90729 human papil	376	5	55.6	550	2	Q9S4M6	Q9s4m6 escherichia
304	5	55.6	466	13	Q9IAX7	Q9iay7 agkistrodon	377	5	55.6	551	2	Q9KFS2	Q9kfs2 streptomyce
305	5	55.6	466	13	Q9IAX6	Q9iay6 agkistrodon	378	5	55.6	551	2	Q9RSN4	Q9rsn4 deinococcus
306	5	55.6	467	4	Q9NXL1	Q9nxl1 homo sapien	379	5	55.6	555	2	O9S612	O9s612 escherichia
307	5	55.6	471	5	O9Y8B0	O9y8b0 drosophila	380	5	55.6	555	2	O9R3X4	O9r3x4 escherichia
308	5	55.6	473	2	O50210	O50210 synechococc	381	5	55.6	555	4	O9H4Z6	O9h4z6 homo sapien
309	5	55.6	475	2	O9X727	O9x727 escherichia	382	5	55.6	556	2	O9S4M3	O9s4m3 escherichia
310	5	55.6	475	2	O9F8R9	O9f8r9 escherichia	383	5	55.6	556	10	Q9M172	Q9m172 arabisdopsis
311	5	55.6	475	2	O9F8R3	O9f8r3 escherichia	384	5	55.6	556	10	Q9M172	Q9m172 arabisdopsis

385	5	55.6	556	10	09LFW2	Q9LW2 arabidopsis	458	5	55.6	739	5	Q9NB06	Q9NB6 trypanosoma
386	5	55.6	557	2	067869	067869 aquilex aeo	459	5	55.6	739	5	Q18221	Q18221 caenorhabdl
387	5	55.6	559	2	Q9RRY4	Q9RRY4 deinococcus	460	5	55.6	748	3	Q12311	Q12311 saccharomyc
388	5	55.6	560	2	Q9S4M5	Q9S4M5 escherichia	461	5	55.6	748	3	Q9P3D8	Q9P3D8 neurospora
389	5	55.6	561	4	Q92952	Q92952 homo sapien	462	5	55.6	750	2	P96256	P96256 mycobacteri
390	5	55.6	567	2	Q33932	Q33932 listeria mo	463	5	55.6	752	2	Q9K3J3	Q9K3J3 streptococc
391	5	55.6	570	2	Q66690	Q66690 aquilex aeo	464	5	55.6	753	5	Q22436	Q22436 caenorhabdl
392	5	55.6	572	2	Q26042	Q26042 penaeus jap	465	5	55.6	753	11	Q9JHW8	Q9JHW8 rattus norv
393	5	55.6	573	3	Q08179	Q08179 saccharomyc	466	5	55.6	757	10	Q9SR14	Q9SR14 arabidopsis
394	5	55.6	579	2	Q51504	Q51504 pseudomonas	467	5	55.6	760	4	Q00199	Q00199 homo sapien
395	5	55.6	581	2	Q9RU45	Q9RU45 deinococcus	468	5	55.6	763	2	Q9ZBL8	Q9ZBL8 mycobacteri
396	5	55.6	583	2	Q9ZNT9	Q9ZNT9 pseudomonas	469	5	55.6	781	10	Q9LWJ6	Q9LWJ6 arabidopsis
397	5	55.6	584	4	Q9UNR7	Q9UNR7 homo sapien	470	5	55.6	782	4	Q9UI94	Q9UI94 homo sapien
398	5	55.6	585	2	P71246	P71246 escherichia	471	5	55.6	784	5	Q9VIS7	Q9VIS7 drosophila
399	5	55.6	587	2	Q9L736	Q9L736 escherichia	472	5	55.6	796	5	Q16145	Q16145 caenorhabdl
400	5	55.6	585	2	Q9K2Y6	Q9K2Y6 escherichia	473	5	55.6	797	4	Q95267	Q95267 homo sapien
401	5	55.6	587	10	Q9MOT2	Q9MOT2 arabidopsis	474	5	55.6	797	4	Q9UNN9	Q9UNN9 homo sapien
402	5	55.6	596	10	Q9LSW2	Q9LSW2 arabidopsis	475	5	55.6	803	2	Q9RBR9	Q9RBR9 pseudomonas
403	5	55.6	599	2	Q9PEU0	Q9PEU0 campylobact	476	5	55.6	816	5	Q9VKJ9	Q9VKJ9 drosophila
404	5	55.6	604	2	Q00509	Q00509 streptomyce	477	5	55.6	817	3	Q07229	Q07229 saccharomyc
405	5	55.6	605	1	Q26511	Q26511 methanobact	478	5	55.6	822	5	Q9GUK3	Q9GUK3 leishmania
406	5	55.6	606	2	Q53561	Q53561 streptococc	479	5	55.6	823	5	Q25343	Q25343 leishmania
407	5	55.6	609	13	Q13027	Q13027 xenopus lae	480	5	55.6	829	5	Q61957	Q61957 caenorhabdl
408	5	55.6	612	4	Q93052	Q93052 homo sapien	481	5	55.6	831	4	Q9NKS6	Q9NKS6 homo sapien
409	5	55.6	613	2	Q9PE86	Q9PE86 xylella fas	482	5	55.6	831	11	Q9EP22	Q9EP22 mus musculu
410	5	55.6	614	3	Q12662	Q12662 phanerocha	483	5	55.6	833	1	Q9S594	Q9S594 mycoplasma
411	5	55.6	620	5	Q9YOD4	Q9YOD4 penaeus mon	484	5	55.6	834	1	Q29217	Q29217 archaeoglob
412	5	55.6	620	5	Q9NEC8	Q9NEC8 leishmania	485	5	55.6	835	5	Q9V744	Q9V744 drosophila
413	5	55.6	624	4	Q9P2U4	Q9P2U4 homo sapien	486	5	55.6	836	4	Q9P2J0	Q9P2J0 homo sapien
414	5	55.6	626	4	Q9P2U3	Q9P2U3 homo sapien	487	5	55.6	842	2	Q52590	Q52590 pseudomonas
415	5	55.6	637	2	P72813	P72813 synechocyst	488	5	55.6	842	5	Q9VGC9	Q9VGC9 drosophila
416	5	55.6	637	5	Q9WDE6	Q9WDE6 drosophila	489	5	55.6	843	2	Q07289	Q07289 streptococc
417	5	55.6	642	2	Q9FC03	Q9FC03 streptomyce	490	5	55.6	846	5	Q9VZ10	Q9VZ10 drosophila
418	5	55.6	648	2	Q9EYC8	Q9EYC8 escherichia	491	5	55.6	847	14	Q41537	Q41537 human immun
419	5	55.6	648	4	Q14760	Q14760 homo sapien	492	5	55.6	849	4	Q9NVE4	Q9NVE4 homo sapien
420	5	55.6	651	4	Q9WZ5	Q9WZ5 homo sapien	493	5	55.6	859	14	Q9YVZ4	Q9YVZ4 human immun
421	5	55.6	651	5	Q18181	Q18181 caenorhabdl	494	5	55.6	864	5	Q9VGC8	Q9VGC8 drosophila
422	5	55.6	653	4	Q9NQT9	Q9NQT9 homo sapien	495	5	55.6	866	4	Q9H466	Q9H466 homo sapien
423	5	55.6	659	5	Q76517	Q76517 trypanosoma	496	5	55.6	870	5	Q9GRY4	Q9GRY4 leishmania
424	5	55.6	660	5	Q9VTE3	Q9VTE3 drosophila	497	5	55.6	872	2	Q9S4J3	Q9S4J3 streptococc
425	5	55.6	662	3	Q94041	Q94041 candida alb	498	5	55.6	875	4	Q60278	Q60278 homo sapien
426	5	55.6	665	10	Q65793	Q65793 pharbitis n	499	5	55.6	882	5	Q91752	Q91752 drosophila
427	5	55.6	666	4	Q9H907	Q9H907 homo sapien	500	5	55.6	883	2	Q9RF11	Q9RF11 myxococcus
428	5	55.6	668	5	Q9VWM7	Q9VWM7 drosophila	501	5	55.6	884	11	Q9JUZ2	Q9JUZ2 rattus norv
429	5	55.6	670	4	Q75714	Q75714 homo sapien	502	5	55.6	899	3	Q06132	Q06132 saccharomyc
430	5	55.6	670	4	Q9UPP5	Q9UPP5 homo sapien	503	5	55.6	920	3	P78953	P78953 schizosacch
431	5	55.6	670	4	Q9UNK7	Q9UNK7 homo sapien	504	5	55.6	929	1	Q9YDS4	Q9YDS4 aeropyrum p
432	5	55.6	670	6	Q9XTE2	Q9XTE2 cercopithe	505	5	55.6	929	10	Q23410	Q23410 arabidopsis
433	5	55.6	672	4	Q9H8W8	Q9H8W8 homo sapien	506	5	55.6	930	11	Q9JUZ3	Q9JUZ3 rattus norv
434	5	55.6	675	10	Q9FUI4	Q9FUI4 arabidopsis	507	5	55.6	930	11	Q9EYF3	Q9EYF3 mus musculu
435	5	55.6	681	4	Q75943	Q75943 homo sapien	508	5	55.6	930	11	Q9EQH2	Q9EQH2 mus musculu
436	5	55.6	682	1	Q9HPZ0	Q9HPZ0 halobacteri	509	5	55.6	931	10	Q04026	Q04026 arabidopsis
437	5	55.6	682	4	Q9NMX3	Q9NMX3 homo sapien	510	5	55.6	935	2	Q9XBP4	Q9XBP4 myxococcus
438	5	55.6	682	10	Q9M3E5	Q9M3E5 arabidopsis	511	5	55.6	939	14	Q9NRN1	Q9NRN1 macaca mula
439	5	55.6	683	10	Q9SK94	Q9SK94 arabidopsis	512	5	55.6	939	14	Q9JZ19	Q9JZ19 macaca mula
440	5	55.6	685	4	Q9NR61	Q9NR61 homo sapien	513	5	55.6	940	10	Q81471	Q81471 arabidopsis
441	5	55.6	685	11	Q9JHZ7	Q9JHZ7 mus musculu	514	5	55.6	941	4	Q9UKY2	Q9UKY2 homo sapien
442	5	55.6	686	11	Q9J171	Q9J171 mus musculu	515	5	55.6	941	4	Q9UHF8	Q9UHF8 homo sapien
443	5	55.6	687	11	Q89024	Q89024 mus musculu	516	5	55.6	941	4	Q9NZ08	Q9NZ08 caenorhabdl
444	5	55.6	687	14	Q88271	Q88271 simian parv	517	5	55.6	964	5	Q10953	Q10953 streptococc
445	5	55.6	688	11	Q88934	Q88934 mus musculu	518	5	55.6	969	2	Q9RZ15	Q9RZ15 deinococcus
446	5	55.6	689	5	Q01719	Q01719 strongyloce	519	5	55.6	969	4	Q9Y214	Q9Y214 homo sapien
447	5	55.6	691	13	Q57658	Q57658 gallus gall	520	5	55.6	976	2	Q9EVI7	Q9EVI7 actinomyces
448	5	55.6	692	2	Q9FZJ1	Q9FZJ1 streptomyce	521	5	55.6	976	2	Q9EVI7	Q9EVI7 actinomyces
449	5	55.6	693	2	Q9K437	Q9K437 streptomyce	522	5	55.6	977	3	Q94540	Q94540 schizosacch
450	5	55.6	694	4	Q9P276	Q9P276 homo sapien	523	5	55.6	977	13	Q91925	Q91925 xenopus lae
451	5	55.6	700	5	Q96839	Q96839 drosophila	524	5	55.6	988	5	Q16851	Q16851 pluteia xy
452	5	55.6	700	5	Q9V887	Q9V887 drosophila	525	5	55.6	990	13	Q91803	Q91803 xenopus lae
453	5	55.6	700	5	Q9UAS6	Q9UAS6 drosophila	526	5	55.6	1004	2	P71719	P71719 mycobacteri
454	5	55.6	708	4	Q9NVU0	Q9NVU0 homo sapien	527	5	55.6	1006	5	Q9W3Y4	Q9W3Y4 drosophila
455	5	55.6	730	10	Q9FF38	Q9FF38 arabidopsis	528	5	55.6	1008	13	Q9DER7	Q9DER7 gallus gall
456	5	55.6	732	5	Q9YOC4	Q9YOC4 leishmania	529	5	55.6	1013	4	Q43897	Q43897 homo sapien
457	5	55.6	735	13	Q57381	Q57381 xenopus lae	530	5	55.6	1013	4	Q9NDS4	Q9NDS4 homo sapien

531	5	55.6	1022	5	061907	061907 caenorhabd	604	5	55.6	1719	11	P97789	P97789 mus musculus
532	5	55.6	1022	13	057460	057460 streptococc	605	5	55.6	1723	11	060767	060767 mus musculus
533	5	55.6	1026	2	09X3R6	09X3R6 streptococc	606	5	55.6	1766	14	060759	060759 fowlpox vir
534	5	55.6	1030	4	014974	014974 homo sapien	607	5	55.6	1769	6	097758	097758 canis famli
535	5	55.6	1047	4	09NSZ0	09NSZ0 homo sapien	608	5	55.6	1777	14	09J5B3	09J5B3 fowlpox vir
536	5	55.6	1057	5	09W294	09W294 drosophila	609	5	55.6	1785	5	025685	025685 plasmodium
537	5	55.6	1066	4	09UGX4	09UGX4 homo sapien	610	5	55.6	1798	5	09VUB6	09VUB6 drosophila
538	5	55.6	1070	5	094290	094290 caenorhabd	611	5	55.6	1799	5	018220	018220 caenorhabd
539	5	55.6	1074	5	09VUZ2	09VUZ2 drosophila	612	5	55.6	1802	5	018219	018219 caenorhabd
540	5	55.6	1103	9	09G0B5	09G0B5 bacterioph	613	5	55.6	1802	14	09J5C2	09J5C2 fowlpox vir
541	5	55.6	1108	3	P87121	P87121 schizosacch	614	5	55.6	1822	2	007290	007290 streptococc
542	5	55.6	1112	10	09FH47	09FH47 arabidops	615	5	55.6	1829	5	019815	019815 caenorhabd
543	5	55.6	1113	2	09K6S8	09K6S8 bacillus ha	616	5	55.6	1846	10	09M0B6	09M0B6 arabidops
544	5	55.6	1115	4	09ULI0	09ULI0 homo sapien	617	5	55.6	1848	5	09VUB6	09VUB6 drosophila
545	5	55.6	1119	3	09P3N5	09P3N5 neurospora	618	5	55.6	1870	14	09J5A0	09J5A0 fowlpox vir
546	5	55.6	1124	5	09NDF3	09NDF3 apis mellif	619	5	55.6	1912	14	09J5C3	09J5C3 fowlpox vir
547	5	55.6	1135	3	09UVU5	09UVU5 pichia angu	620	5	55.6	1963	10	09LX79	09LX79 arabidops
548	5	55.6	1140	5	09VTF5	09VTF5 drosophila	621	5	55.6	1967	4	09P2D1	09P2D1 homo sapien
549	5	55.6	1159	4	09UF39	09UF39 homo sapien	622	5	55.6	1983	5	09U165	09U165 leishmania
550	5	55.6	1175	6	09TV66	09TV66 oryctolagus	623	5	55.6	2001	10	09M659	09M659 arabidops
551	5	55.6	1182	14	002711	002711 murine endo	624	5	55.6	2001	10	09M658	09M658 arabidops
552	5	55.6	1187	5	09V7N3	09V7N3 drosophila	625	5	55.6	2164	13	09IAR9	09IAR9 galus gall
553	5	55.6	1188	4	09UQ26	09UQ26 homo sapien	626	5	55.6	2210	5	0241B9	0241B9 drosophila
554	5	55.6	1191	5	09VJN7	09VJN7 drosophila	627	5	55.6	2217	5	09VW15	09VW15 drosophila
555	5	55.6	1199	6	028139	028139 bos taurus	628	5	55.6	2254	10	09LND2	09LND2 arabidops
556	5	55.6	1201	5	09W1J0	09W1J0 drosophila	629	5	55.6	2271	5	09W1A9	09W1A9 drosophila
557	5	55.6	1208	2	09XD11	09XD11 legionella	630	5	55.6	2271	13	09I909	09I909 xenopus lae
558	5	55.6	1211	11	035223	035223 mus musculu	631	5	55.6	2280	5	09V8E6	09V8E6 drosophila
559	5	55.6	1212	2	09L1C8	09L1C8 streptomyce	632	5	55.6	2287	5	019761	019761 caenorhabd
560	5	55.6	1217	5	P91457	P91457 caenorhabd	633	5	55.6	2292	14	089272	089272 encephalomy
561	5	55.6	1217	11	088573	088573 mus musculu	634	5	55.6	2292	14	066763	066763 encephalomy
562	5	55.6	1219	5	09VTF6	09VTF6 drosophila	635	5	55.6	2292	14	066765	066765 encephalomy
563	5	55.6	1224	3	013637	013637 schizosacch	636	5	55.6	2292	14	066850	066850 encephalomy
564	5	55.6	1228	3	012754	012754 saccharomyc	637	5	55.6	2293	14	083422	083422 mengo encp
565	5	55.6	1240	10	09T0B7	09T0B7 arabidops	638	5	55.6	2302	5	09N693	09N693 drosophila
566	5	55.6	1246	4	015052	015052 homo sapien	639	5	55.6	2310	5	09G9A9	09G9A9 drosophila
567	5	55.6	1254	4	09UBS9	09UBS9 homo sapien	640	5	55.6	2344	5	09N3Y8	09N3Y8 caenorhabd
568	5	55.6	1255	5	09NKK8	09NKK8 drosophila	641	5	55.6	2382	4	09E4A3	09E4A3 homo sapien
569	5	55.6	1262	5	09N911	09N911 leishmania	642	5	55.6	2404	5	09VEE3	09VEE3 drosophila
570	5	55.6	1268	11	063623	063623 rattus norv	643	5	55.6	2446	4	09NS05	09NS05 homo sapien
571	5	55.6	1270	4	09HCG5	09HCG5 homo sapien	644	5	55.6	2447	5	022463	022463 caenorhabd
572	5	55.6	1276	11	P97260	P97260 cricetus	645	5	55.6	2500	4	002646	002646 homo sapien
573	5	55.6	1277	4	012770	012770 homo sapien	646	5	55.6	2768	4	09VCO0	09VCO0 drosophila
574	5	55.6	1300	11	09WTL4	09WTL4 mus musculu	647	5	55.6	2785	4	075691	075691 homo sapien
575	5	55.6	1322	5	09VYH2	09VYH2 drosophila	648	5	55.6	2796	2	048926	048926 mycobacteri
576	5	55.6	1331	1	09YEG3	09YEG3 aeropyrum p	649	5	55.6	2836	5	09NKP7	09NKP7 leishmania
577	5	55.6	1341	4	09UFP8	09UFP8 homo sapien	650	5	55.6	2977	14	041892	041892 hepatitis g
578	5	55.6	1354	5	09VG11	09VG11 drosophila	651	5	55.6	2977	5	09VAP9	09VAP9 drosophila
579	5	55.6	1357	5	09W4M4	09W4M4 drosophila	652	5	55.6	3016	14	092531	092531 hepatitis c
580	5	55.6	1400	11	09ESU6	09ESU6 mus musculu	653	5	55.6	3131	3	000869	000869 fusarium eq
581	5	55.6	1405	4	09ULI04	09ULI04 homo sapien	654	5	55.6	3261	4	09Y556	09Y556 homo sapien
582	5	55.6	1415	5	061442	061442 caenorhabd	655	5	55.6	3418	4	013879	013879 homo sapien
583	5	55.6	1418	5	09U6B9	09U6B9 caenorhabd	656	5	55.6	3550	11	09UJN2	09UJN2 mus musculu
584	5	55.6	1423	5	09W1A0	09W1A0 drosophila	657	5	55.6	3576	11	09Q2W2	09Q2W2 mus musculu
585	5	55.6	1431	5	09U6R1	09U6R1 caenorhabd	658	5	55.6	4162	13	098918	098918 galus gall
586	5	55.6	1449	5	010922	010922 caenorhabd	659	5	55.6	4833	11	09QXY6	09QXY6 mus musculu
587	5	55.6	1457	5	09WY35	09WY35 drosophila	660	5	55.6	4857	4	014687	014687 homo sapien
588	5	55.6	1490	5	09U4B5	09U4B5 drosophila	661	5	55.6	5038	11	09QIX7	09QIX7 mus musculu
589	5	55.6	1504	13	09PTW9	09PTW9 opsanus bet	662	5	55.6	5262	4	014686	014686 homo sapien
590	5	55.6	1534	6	028298	028298 canis famli	663	5	55.6	5388	5	09U1D0	09U1D0 leishmania
591	5	55.6	1543	5	09VWAO	09VWAO drosophila	664	5	55.6	6831	5	023550	023551 caenorhabd
592	5	55.6	1566	2	053414	053414 streptococc	665	5	55.6	7160	5	023551	023551 caenorhabd
593	5	55.6	1566	2	055308	055308 streptococc	666	5	55.6	7160	10	09J1G8	09J1G8 tt virus
594	5	55.6	1569	2	054183	054183 streptococc	667	5	55.6	7160	10	09J1G8	09J1G8 tt virus
595	5	55.6	1576	13	09I908	09I908 xenopus lae	668	5	55.6	7160	22	030702	030702 chlamydia t
596	5	55.6	1616	10	09M1Z7	09M1Z7 arabidops	669	5	55.6	7160	23	09R410	09R410 bordetella
597	5	55.6	1624	5	09U679	09U679 strongyloce	670	5	55.6	7160	25	063334	063334 rattus norv
598	5	55.6	1663	4	09U001	09U001 homo sapien	671	5	55.6	7160	25	09J1I1	09J1I1 tt virus
599	5	55.6	1687	11	035651	035651 mus musculu	672	5	55.6	7160	26	09M0Y5	09M0Y5 homo sapien
600	5	55.6	1688	5	09Y9D0	09Y9D0 drosophila	673	5	55.6	7160	26	09M0Y5	09M0Y5 tt virus
601	5	55.6	1694	5	09I4N8	09I4N8 pseudomonas	674	5	55.6	7160	26	09M0Y5	09M0Y5 tt virus
602	5	55.6	1706	11	P97790	P97790 mus musculu	675	5	55.6	7160	26	09M0Y5	09M0Y5 tt virus
603	5	55.6	1711	2	P96311	P96311 anaerocellu	676	5	55.6	7160	26	09M0Y5	09M0Y5 tt virus

677	4	44.4	26	14	09MNU0	09wnj0	tt	virus.	p	750	4	44.4	40	7	09TN24	09tn24	rattus	norv
678	4	44.4	26	14	09W872	09w872	tt	virus.	p	751	4	44.4	41	7	09TN25	09tn25	rattus	norv
679	4	44.4	26	14	09J116	09j116	tt	virus.	o	752	4	44.4	43	2	09PE48	09pe48	xylella	fas
680	4	44.4	26	14	09J115	09j115	tt	virus.	o	753	4	44.4	43	7	09TN26	09tn26	rattus	norv
681	4	44.4	26	14	09J114	09j114	tt	virus.	o	754	4	44.4	44	5	061051	061051	trypnosoma	
682	4	44.4	26	14	09J113	09j113	tt	virus.	o	755	4	44.4	44	7	09S562	09s562	rattus	norv
683	4	44.4	26	14	09J112	09j112	tt	virus.	o	756	4	44.4	45	2	086765	086765	streptomyce	
684	4	44.4	26	14	09J110	09j110	tt	virus.	o	757	4	44.4	47	14	09YR02	09yr02	tt	virus. h
685	4	44.4	26	14	09J1H9	09j1H9	tt	virus.	o	758	4	44.4	47	14	09W7S4	09w7s4	tt	virus. h
686	4	44.4	26	14	09J1H8	09j1H8	tt	virus.	o	759	4	44.4	47	14	09PX67	09px67	watermelon	
687	4	44.4	26	14	09J1H3	09j1H3	tt	virus.	o	760	4	44.4	47	14	09DHM3	09dhm3	watermelon	
688	4	44.4	26	14	09J1H2	09j1H2	tt	virus.	o	761	4	44.4	49	2	045823	045823	quercus	quercus
689	4	44.4	26	14	09J1H0	09j1H0	tt	virus.	o	762	4	44.4	49	6	09M2L9	09m2l9	canis	famil
690	4	44.4	26	14	09J1G6	09j1G6	tt	virus.	o	763	4	44.4	49	14	09JGCT3	09jgct3	tt	virus. p
691	4	44.4	26	14	09J1G5	09j1G5	tt	virus.	o	764	4	44.4	49	14	09JGCT0	09jgct0	tt	virus. p
692	4	44.4	26	14	09J1G4	09j1G4	tt	virus.	o	765	4	44.4	49	14	09DYC0	09dyc0	tt	virus. h
693	4	44.4	26	14	09J1G1	09j1G1	tt	virus.	o	766	4	44.4	51	5	017090	017090	ascaris	suu
694	4	44.4	26	14	09J1G0	09j1G0	tt	virus.	o	767	4	44.4	51	6	018724	018724	hylobates	s
695	4	44.4	26	14	09J1F9	09j1F9	tt	virus.	o	768	4	44.4	51	6	018725	018725	hylobates	s
696	4	44.4	26	14	09J1F8	09j1F8	tt	virus.	o	769	4	44.4	52	2	052864	052864	rhizobium	1
697	4	44.4	26	14	09J1F7	09j1F7	tt	virus.	o	770	4	44.4	52	8	0922U6	0922u6	littorina	s
698	4	44.4	26	14	09J1F6	09j1F6	tt	virus.	o	771	4	44.4	53	1	09YC81	09yc81	aeropyrum	p
699	4	44.4	26	14	09J1F5	09j1F5	tt	virus.	o	772	4	44.4	53	2	09RE85	09re85	staphylococ	
700	4	44.4	26	14	09J1F3	09j1F3	tt	virus.	o	773	4	44.4	53	8	09GB33	09gb33	quiscalus	q
701	4	44.4	26	14	09J1F2	09j1F2	tt	virus.	o	774	4	44.4	54	14	09E3W1	09e3w1	tomato	yell
702	4	44.4	26	14	09J1F1	09j1F1	tt	virus.	o	775	4	44.4	55	5	09ND27	09nd27	leishmania	
703	4	44.4	26	14	09J1F0	09j1F0	tt	virus.	o	776	4	44.4	55	6	018722	018722	pan	troglod
704	4	44.4	26	14	09J1E8	09j1E8	tt	virus.	o	777	4	44.4	55	8	09T3G7	09t3g7	icterus	obe
705	4	44.4	26	14	09J1E7	09j1E7	tt	virus.	o	778	4	44.4	55	8	09MDP1	09mdp1	quiscalus	1
706	4	44.4	26	14	09J1E6	09j1E6	tt	virus.	o	779	4	44.4	55	8	09MDP9	09mdp9	quiscalus	1
707	4	44.4	26	14	09J1E4	09j1E4	tt	virus.	o	780	4	44.4	55	11	063761	063761	rattus	norv
708	4	44.4	26	14	09J1E3	09j1E3	tt	virus.	o	781	4	44.4	57	6	028471	028471	macaca	fasc
709	4	44.4	26	14	09J1E1	09j1E1	tt	virus.	o	782	4	44.4	57	14	09E4B6	09e4b6	human	adeno
710	4	44.4	26	14	09J1E0	09j1E0	tt	virus.	o	783	4	44.4	58	2	09S4C7	09s4c7	poirphyromon	
711	4	44.4	26	14	09J1D9	09j1D9	tt	virus.	o	784	4	44.4	59	14	003812	003812	human	immun
712	4	44.4	26	14	09J1D6	09j1D6	tt	virus.	o	785	4	44.4	60	2	086035	086035	rhizobium	m
713	4	44.4	26	14	09J1D4	09j1D4	tt	virus.	o	786	4	44.4	60	4	014889	014889	homo	sepien
714	4	44.4	26	14	09J1D3	09j1D3	tt	virus.	o	787	4	44.4	61	4	09N0B8	09n0b8	homo	sepien
715	4	44.4	26	14	09J1D2	09j1D2	tt	virus.	o	788	4	44.4	61	14	P89262	P89262	xestia	c-n1
716	4	44.4	26	14	09J1D1	09j1D1	tt	virus.	o	789	4	44.4	62	5	09XWL6	09xwl6	caenorhabdi	
717	4	44.4	26	14	09J1D0	09j1D0	tt	virus.	o	790	4	44.4	62	6	09XS88	09xs88	equus	cabal
718	4	44.4	28	14	09WNI9	09wni9	tt	virus.	o	791	4	44.4	63	5	09NME5	09nme5	leishmania	
719	4	44.4	28	14	09WNI7	09wni7	tt	virus.	o	792	4	44.4	63	10	09XGK9	09xgk9	euglena	gra
720	4	44.4	28	14	09WNI4	09wni4	tt	virus.	o	793	4	44.4	64	2	057026	057026	borrelia	af
721	4	44.4	28	14	09WNI3	09wni3	tt	virus.	o	794	4	44.4	65	5	09GXM2	09gxm2	leishmania	
722	4	44.4	28	14	09QUT7	09qt7	tt	virus.	o	795	4	44.4	66	5	09NE55	09ne55	dirosophila	
723	4	44.4	28	14	09QUT6	09qt6	tt	virus.	o	796	4	44.4	66	7	029916	029916	homo	sepien
724	4	44.4	28	14	09QUT5	09qt5	tt	virus.	o	797	4	44.4	66	8	096802	096802	skeletonema	
725	4	44.4	28	14	09QUT3	09qt3	tt	virus.	o	798	4	44.4	67	2	09WYU9	09wyu9	thermotoga	
726	4	44.4	29	14	09QUT2	09qt2	tt	virus.	o	799	4	44.4	68	2	048848	048848	lactobacilli	
727	4	44.4	29	14	086257	086257	human	rotav	o	800	4	44.4	68	2	09KZG1	09kzg1	streptomyce	
728	4	44.4	29	14	086259	086259	human	rotav	o	801	4	44.4	68	14	09YK4	09yk4	tomato	yell
729	4	44.4	29	14	086262	086262	human	rotav	o	802	4	44.4	69	2	09RCV0	09rcv0	streptomyce	
730	4	44.4	29	14	086256	086256	human	rotav	o	803	4	44.4	69	2	09KMY2	09kmy2	vibrio	chol
731	4	44.4	29	14	086260	086260	human	rotav	o	804	4	44.4	69	11	09JMB7	09jmb7	mus	musculu
732	4	44.4	29	14	086261	086261	human	rotav	o	805	4	44.4	69	14	P89655	P89655	murid	herpe
733	4	44.4	29	14	086263	086263	human	rotav	o	806	4	44.4	69	14	09PE27	09pe27	simlan	herp
734	4	44.4	29	14	086264	086264	human	rotav	o	807	4	44.4	70	14	089817	089817	borna	disea
735	4	44.4	29	14	086265	086265	human	rotav	o	808	4	44.4	71	2	09PAB8	09pab8	xylella	fas
736	4	44.4	29	14	086266	086266	human	rotav	o	809	4	44.4	71	10	09PTW6	09ptw6	oryza	saliv
737	4	44.4	29	14	086267	086267	human	rotav	o	810	4	44.4	71	14	087616	087616	chimpanzee	
738	4	44.4	31	14	09IMY6	09imy6	tt	virus.	h	811	4	44.4	72	10	09LX10	09lx10	arabidopsi	s
739	4	44.4	31	14	09IMY5	09imy5	tt	virus.	h	812	4	44.4	72	14	085982	085982	human	immun
740	4	44.4	31	14	09IMY4	09imy4	tt	virus.	h	813	4	44.4	73	1	027331	027331	methanobact	
741	4	44.4	31	14	09IMY2	09imy2	tt	virus.	h	814	4	44.4	73	2	09PES8	09pes8	xylella	fas
742	4	44.4	32	2	047657	047657	escherichia		h	815	4	44.4	73	4	09KX54	09kx54	streptomyce	
743	4	44.4	32	14	09WNI5	09wni5	tt	virus.	o	816	4	44.4	73	5	09X039	09x039	homo	sepien
744	4	44.4	32	14	09JG31	09jg31	tt	virus.	o	817	4	44.4	73	5	09X039	09x039	caenorhabdi	
745	4	44.4	32	14	09JG27	09jg27	tt	virus.	o	818	4	44.4	74	12	P87655	P87655	europian	br
746	4	44.4	35	14	09YK29	09yk29	human	immun	o	819	4	44.4	74	14	092HMB	092hmb	bruceella	me
747	4	44.4	37	11	060704	060704	mus	musculu	o	820	4	44.4	74	14	056576	056576	hepatitis	c
748	4	44.4	38	8	09T2N6	09t2n6	spina	cia ol	o	821	4	44.4	74	14	081723	081723	hepatitis	c
749	4	44.4	39	10	040959	040959	physcomlitr		o	822	4	44.4	74	14	081743	081743	hepatitis	c

823	4	44.4	74	14	081744	081744 hepatitis c	896	4	44.4	85	14	070695	070695 human immun
824	4	44.4	74	14	081745	081745 hepatitis c	897	4	44.4	86	14	090298	090298 odocolleus
825	4	44.4	74	14	09YP05	09YP05 hepatitis c	898	4	44.4	85	14	049372	049372 mycoplasma
826	4	44.4	74	14	09YP04	09YP04 hepatitis c	899	4	44.4	86	2	P94196	P94196 agrobacteri
827	4	44.4	74	14	09YP03	09YP03 hepatitis c	900	4	44.4	86	6	09NV23	09NV23 oryctolagus
828	4	44.4	74	14	09YP02	09YP02 hepatitis c	901	4	44.4	86	6	P82672	P82672 bos taurus
829	4	44.4	74	14	09YPP3	09YPP3 hepatitis c	902	4	44.4	86	10	042283	042283 arabidopsis
830	4	44.4	74	14	09YPP1	09YPP1 hepatitis c	903	4	44.4	87	2	09L8H3	09L8H3 streptomyce
831	4	44.4	74	14	09YJB8	09YJB8 hepatitis c	904	4	44.4	87	5	09W216	09W216 drosophila
832	4	44.4	74	14	09YJAS	09YJAS hepatitis c	905	4	44.4	87	5	09WAC9	09WAC9 caenorhabdi
833	4	44.4	74	14	09YJDO	09YJDO hepatitis c	906	4	44.4	87	5	09GWT4	09GWT4 leishmania
834	4	44.4	74	14	09WK43	09WK43 hepatitis c	907	4	44.4	88	2	09RGW7	09RGW7 actinobact
835	4	44.4	74	14	09WK42	09WK42 hepatitis c	908	4	44.4	88	10	Q41185	Q41185 arabidopsis
836	4	44.4	75	2	044892	044892 borrelia bu	909	4	44.4	88	14	09Q0P9	09Q0P9 hepatitis c
837	4	44.4	75	2	09TX66	09TX66 drosophila	910	4	44.4	88	14	09Q0P9	09Q0P9 hepatitis c
838	4	44.4	75	14	069312	069312 gallid herp	911	4	44.4	88	14	09Q0P4	09Q0P4 hepatitis c
839	4	44.4	76	2	045627	045627 bacillus st	912	4	44.4	89	2	09ZNI3	09ZNI3 pseudomonas
840	4	44.4	76	2	09RS66	09RS66 delnoccocus	913	4	44.4	89	2	09LOV2	09LOV2 streptomyce
841	4	44.4	76	2	09KGR1	09KGR1 vibrio para	914	4	44.4	89	2	030687	030687 rhodococcus
842	4	44.4	76	9	09MBW6	09MBW6 bacterioph	915	4	44.4	91	2	09Z982	09Z982 chlamydia p
843	4	44.4	76	10	022492	022492 oenanthel ja	916	4	44.4	91	2	09ZTSJ1	09ZTSJ1 chlamydia p
844	4	44.4	76	10	082581	082581 pimpinella	917	4	44.4	91	5	09N3U7	09N3U7 caenorhabdi
845	4	44.4	76	10	09XIK9	09XIK9 arabidopsis	918	4	44.4	91	5	09G6W3	09G6W3 ostercagia
846	4	44.4	76	14	084503	084503 paramecium	919	4	44.4	92	2	09ZBZ9	09ZBZ9 streptomyce
847	4	44.4	77	5	09PBY2	09PBY2 xylella fas	920	4	44.4	92	10	09LC69	09LC69 oryza sativ
848	4	44.4	77	5	09N3U6	09N3U6 caenorhabdi	921	4	44.4	93	2	052948	052948 bacillus su
849	4	44.4	77	14	09QPC7	09QPC7 tomato yell	922	4	44.4	93	2	09Z9G9	09Z9G9 chlamydia p
850	4	44.4	78	2	086789	086789 streptomyce	923	4	44.4	94	6	09N030	09N030 macaca fasc
851	4	44.4	78	2	034155	034155 microcystis	924	4	44.4	94	10	09SIN7	09SIN7 arabidopsis
852	4	44.4	78	2	049654	049654 mycobacteri	925	4	44.4	95	14	005535	005535 equine herp
853	4	44.4	78	7	095413	095413 ovis aries	926	4	44.4	96	2	033423	033423 pseudomonas
854	4	44.4	79	2	047383	047383 escherichia	927	4	44.4	96	5	019828	019828 caenorhabdi
855	4	44.4	79	5	017013	017013 anopheles g	928	4	44.4	96	5	09VZB7	09VZB7 drosophila
856	4	44.4	79	5	094188	094188 caenorhabdi	929	4	44.4	96	5	09GWE7	09GWE7 leishmania
857	4	44.4	79	5	09W7T5	09W7T5 leishmania	930	4	44.4	96	10	09LJU3	09LJU3 arabidopsis
858	4	44.4	80	10	09ZT55	09ZT55 zea mays (m	931	4	44.4	97	2	09ZNP5	09ZNP5 comamonas t
859	4	44.4	80	10	09M697	09M697 brassica ol	932	4	44.4	97	5	09VHF4	09VHF4 drosophila
860	4	44.4	81	5	094209	094209 caenorhabdi	933	4	44.4	98	2	09WX46	09WX46 escherichia
861	4	44.4	81	5	09NMKO	09NMKO leishmania	934	4	44.4	98	5	09N616	09N616 drosophila
862	4	44.4	81	6	028856	028856 sus scrofa	935	4	44.4	98	11	061369	061369 mus musclu
863	4	44.4	82	2	P95312	P95312 mycobacteri	936	4	44.4	99	1	09UX22	09UX22 sulfobolus
864	4	44.4	82	2	033259	033259 escherichia	937	4	44.4	99	2	044464	044464 agrobacteri
865	4	44.4	82	5	094210	094210 caenorhabdi	938	4	44.4	99	14	09E586	09E586 human immun
866	4	44.4	82	5	094212	094212 caenorhabdi	939	4	44.4	100	1	09YBJ1	09YBJ1 aeropyrum p
867	4	44.4	82	5	094214	094214 caenorhabdi	940	4	44.4	100	1	09UWS1	09UWS1 sulfobolus
868	4	44.4	82	5	061784	061784 caenorhabdi	941	4	44.4	100	5	P92194	P92194 drosophila
869	4	44.4	83	2	09WTC3	09WTC3 escherichia	942	4	44.4	100	14	040501	040501 human immun
870	4	44.4	83	7	019724	019724 homo sapien	943	4	44.4	100	14	069359	069359 feline herp
871	4	44.4	83	10	09FW11	09FW11 oryza sativ	944	4	44.4	100	14	09J7Y9	09J7Y9 feline aden
872	4	44.4	84	5	016803	016803 paracentro	945	4	44.4	100	14	09DIZ8	09DIZ8 human adeno
873	4	44.4	84	10	024072	024072 medicago sa	946	4	44.4	100	14	09DIZ7	09DIZ7 human adeno
874	4	44.4	84	14	090694	090694 unidentified	947	4	44.4	100	14	09DIZ6	09DIZ6 human adeno
875	4	44.4	84	14	092612	092612 unidentified	948	4	44.4	100	14	09DIZ5	09DIZ5 human adeno
876	4	44.4	84	14	092613	092613 unidentified	949	4	44.4	100	14	09DIZ4	09DIZ4 human adeno
877	4	44.4	84	14	09QPU4	09QPU4 human adeno	950	4	44.4	100	14	09DIZ3	09DIZ3 human adeno
878	4	44.4	84	14	09QPU3	09QPU3 human adeno	951	4	44.4	100	14	09DIZ2	09DIZ2 human adeno
879	4	44.4	84	14	09QPU2	09QPU2 human adeno	952	4	44.4	100	14	09DIZ1	09DIZ1 human adeno
880	4	44.4	84	14	09QPU1	09QPU1 human adeno	953	4	44.4	100	14	09DIZ0	09DIZ0 human adeno
881	4	44.4	84	14	09QPU0	09QPU0 human adeno	954	4	44.4	100	14	09DIT9	09DIT9 human adeno
882	4	44.4	84	14	09QPT9	09QPT9 human adeno	955	4	44.4	100	14	09DIT8	09DIT8 human adeno
883	4	44.4	84	14	09QPT8	09QPT8 human adeno	956	4	44.4	100	14	09DIT7	09DIT7 human adeno
884	4	44.4	84	14	09QPT7	09QPT7 human adeno	957	4	44.4	100	14	09DIT6	09DIT6 human adeno
885	4	44.4	84	14	09QPT6	09QPT6 human adeno	958	4	44.4	101	2	09Z8B1	09Z8B1 chlamydia p
886	4	44.4	84	14	09QPT5	09QPT5 human adeno	959	4	44.4	101	5	017614	017614 caenorhabdi
887	4	44.4	84	14	09QPT4	09QPT4 human adeno	960	4	44.4	101	5	09OIT3	09OIT3 caenorhabdi
888	4	44.4	84	14	09QPT3	09QPT3 human adeno	961	4	44.4	101	14	065385	065385 unidentified
889	4	44.4	84	14	09QPT2	09QPT2 human adeno	962	4	44.4	101	14	065432	065432 unidentified
890	4	44.4	84	14	09QPT1	09QPT1 human adeno	963	4	44.4	101	14	065659	065659 unidentified
891	4	44.4	84	14	09QPT0	09QPT0 human adeno	964	4	44.4	101	14	065679	065679 unidentified
892	4	44.4	84	14	09QPS9	09QPS9 human adeno	965	4	44.4	102	5	09PK42	09PK42 chlamydia m
893	4	44.4	84	14	09QPS8	09QPS8 human adeno	966	4	44.4	102	5	0917C6	0917C6 drosophila
894	4	44.4	84	14	09QPS7	09QPS7 human adeno	967	4	44.4	102	10	004341	004341 arabidopsis
895	4	44.4	85	2	09Z432	09Z432 pseudomonas	968	4	44.4	102	10	P93057	P93057 bromus iner

969 4 44.4 102 10 Q9FEZ8  
970 4 44.4 103 2 Q9RX39  
971 4 44.4 103 5 P90571  
972 4 44.4 103 5 P90572  
973 4 44.4 103 14 Q71233  
974 4 44.4 103 14 Q71244  
975 4 44.4 103 14 Q71226  
976 4 44.4 103 14 Q71231  
977 4 44.4 103 14 P88761  
978 4 44.4 103 14 P88762  
979 4 44.4 103 14 P88763  
980 4 44.4 103 14 Q39135  
981 4 44.4 103 14 Q39137  
982 4 44.4 103 14 Q71228  
983 4 44.4 103 14 Q9MFW8  
984 4 44.4 104 4 Q9P0E3  
985 4 44.4 104 5 P92184  
986 4 44.4 104 5 Q96905  
987 4 44.4 104 5 P92192  
988 4 44.4 105 2 Q54742  
989 4 44.4 105 2 Q87035  
990 4 44.4 105 2 Q9X8C6  
991 4 44.4 105 2 Q9PEK4  
992 4 44.4 105 4 P78449  
993 4 44.4 105 5 Q17410  
994 4 44.4 105 5 P92201  
995 4 44.4 105 5 Q91706  
996 4 44.4 105 8 Q32712  
997 4 44.4 105 8 Q9MG34  
998 4 44.4 105 9 Q80231  
999 4 44.4 105 10 Q41031  
1000 4 44.4 105 11 Q9R1P8

## ALIGNMENTS

Q9FEZ8 glycine max  
Q9RX39 delnoccocus  
P90571 plasmodium  
P90572 plasmodium  
Q71233 human immun  
Q71244 human immun  
Q71226 human immun  
Q71231 human immun  
P88761 human immun  
P88762 human immun  
P88763 human immun  
Q39135 human immun  
Q39137 human immun  
Q71228 human immun  
Q9MFW8 human immun  
Q9P0E3 homo sapien  
P92184 drosophila  
Q96905 drosophila  
P92192 drosophila  
Q54742 streptococc  
Q87035 vibrio chol  
Q9X8C6 streptomyce  
Q9PEK4 xyella fas  
P78449 homo sapien  
Q17410 lucilia cup  
P92201 drosophila  
Q91706 drosophila  
Q32712 nicotiana t  
Q9MG34 actinomyx ju  
Q80231 bacterioph  
Q41031 picea slitch  
Q9R1P8 mus muscula

RESULT 1  
Q9RNU3 PRELIMINARY; PRT; 86 AA.  
ID Q9RNU3  
AC Q9RNU3  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE P60 PROTEIN (FRAGMENT).  
GN IAP.  
OS Listeria monocytogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Listeria.  
OX NCBI\_TaxID=1639;  
RN (1)  
RX SEQUENCE FROM N.A.  
RP MEDLINE=90256283; Pubmed=2111287;  
RA Kohler S., Leimister-Wachter M., Chakraborty T., Lottspeich F.,  
RA Goebel W.;  
RT "The gene coding for protein p60 of Listeria monocytogenes and its use  
RT as a specific probe for Listeria monocytogenes.";  
RL Infect. Immun. 58:1943-1950(1990).  
RN (2)  
RP SEQUENCE FROM N.A.  
RA Park S.;  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF179003; AAD5089.1; -.  
FT NON\_TER  
FT NON\_TER  
SQ SEQUENCE 86 AA; 9194 MM; 71F649A817D697F6 CRC64;

Query Match 100.0%; Score 9; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 0 0046;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTPVAPRO 9

Db 29 VSTPVAPRO 37

RESULT 2  
ID Q03493 PRELIMINARY; PRT; 478 AA.  
AC Q03493  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE PROTEIN P60 PRECURSOR (INVASION-ASSOCIATED PROTEIN).  
GN IAP.  
OS Listeria monocytogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Listeria.  
OX NCBI\_TaxID=1639;  
RN (1)  
RX SEQUENCE FROM N.A.  
RP MEDLINE=93094153; Pubmed=1459966;  
RA Buber A., Kuhn M., Goebel W., Koehler S.;  
RT "Structural and functional properties of the p60 proteins from  
RT different Listeria species.";  
RL J. Bacteriol. 174:8166-8171(1992).  
CC -1- FUNCTION: THIS MAJOR EXTRACELLULAR PROTEIN MAY BE INVOLVED IN THE  
CC INVASION OF NONPROFESSIONAL PHAGOCYTIC CELLS BY LISTERIA.  
DR EMBL; M80351; AAA25280.1; -.  
DR InterPro: IPR000064; -.  
DR InterPro: IPR002482; -.  
DR Pfam: PF00877; NUPC\_P60; 1.  
DR Pfam: PF01476; LysM; 2.  
DR SMART; SM00257; LysM; 1.  
KW SIGNAL.  
FT SIGNAL  
FT CHAIN  
FT CHAIN  
SQ SEQUENCE 478 AA; 49997 MM; D46D050507402344 CRC64;

Query Match 100.0%; Score 9; DB 2; Length 478;  
Best Local Similarity 100.0%; Pred. No. 0 018;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTPVAPRO 9  
Db 145 VSTPVAPRO 153

RESULT 3  
ID Q9DEV2 PRELIMINARY; PRT; 530 AA.  
AC Q9DEV2  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE TARGET OF JUN 3.  
GN TOJ3.  
OS Coturnix coturnix (common quail).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Coturnix.  
OX NCBI\_TaxID=9091;  
RN (1)  
RX SEQUENCE FROM N.A.  
RP Bader A.G., Hartl M., Bister K.;  
RT "TOJ3, a novel potential target of the oncogenic transcription factor  
RT Jun.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY007310; AAG16624.1; -.  
SQ SEQUENCE 530 AA; 57929 MM; E460DB89D9FF490B CRC64;

Query Match 77.8%; Score 7; DB 13; Length 530;  
Best Local Similarity 100.0%; Pred. No. 3;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVP 7  
11111111

DB 169 VSTPVP 175

RESULT 4  
O9HYT7 PRELIMINARY; PRT; 102 AA.

ID O9HYT7  
AC O9HYT7  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL PROTEIN PA3307.  
GN PA3307.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RA MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey R.L., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltz L., Tolentino E., Westbrook-Wadman S., Yan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reiser J., Salter M.H., Hancock R.E.W., Lory S., Olson M.V.,  
RT "Complete genomic sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen";  
RL Nature 406:959-964(2000).  
DR EMBL: AE004733; AAG0695.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 102 AA; 11114 MW; D92FEE48B95286B CRC64;

Query Match 66.7%; Score 6; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPQ 9  
111111

DB 27 PVAPQ 32

RESULT 5  
O9VX73 PRELIMINARY; PRT; 209 AA.

ID O9VX73  
AC O9VX73  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE CG5070.0 PROTEIN.  
GN CG5070.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RA MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Adayant A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dletz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Modarity C., Morris J., Moshireli A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleib J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Splier E., Spindling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003504; AAF48705.1; -.  
DR FlyBase: FBgn0030824; CG5070.  
SQ SEQUENCE 209 AA; 19129 MW; E9F5084CDE0E638A CRC64;

Query Match 66.7%; Score 6; DB 5; Length 209;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVA 6  
111111

DB 135 VSTPVA 140

RESULT 6  
O9FU39 PRELIMINARY; PRT; 237 AA.

ID O9FU39  
AC O9FU39  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE P0001B06.23 PROTEIN.  
GN P0001B06.23.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhacroidae; Oryzae;  
OC Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
RT clone: P0001B06.";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AP002537; BAB16870.1; -.  
SQ SEQUENCE 237 AA; 25810 MW; 972E4FC57C2E39A CRC64;

Query Match 66.7%; Score 6; DB 10; Length 237;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY      3  TPVAPT 8
        |||||
Db      220  TPVAPT 225

RESULT  7
ID      084512      PRELIMINARY;      PRT;      288 AA.
AC      084512;
DT      01-NOV-1998 (TREMBLrel. 08, Created)
DT      01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT      01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE      HYPOTHEICAL 32.2 KDA PROTEIN.
GN      C7504.
OS      Chlamydia trachomatis.
OC      Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX      NCBI_TaxID=813;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=D/UM-3/CX;
RA      Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA      Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA      Davis R.W.;
RT      "Genome sequence of an obligate intracellular pathogen of humans:
RT      Chlamydia trachomatis."
RL      Science 282:754-759(1998).
DR      EMBL; AE001323; AAC68105.1; -.
KW      Hypothetical protein.
SQ      SEQUENCE 288 AA; 32191 MW; 9B5CAE26D04028B5 CRC64;

Query Match      66.7%; Score 6; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2  STPVAP 7
        |||||
Db      131  STPVAP 136

RESULT  8
ID      09KYN4      PRELIMINARY;      PRT;      289 AA.
AC      09KYN4;
DT      01-OCT-2000 (TREMBLrel. 15, Created)
DT      01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT      01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE      PUTATIVE ARAC-FAMILY TRANSCRIPTIONAL REGULATOR.
GN      SC9H11.05.
OS      Streptomyces coelicolor.
OC      Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae;
OC      Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX      NCBI_TaxID=1902;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-A3(2);
RA      Seeger K.J., Harris D.;
RL      Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN-A3(2);
RA      Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL      Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN-A3(2);
RA      MEDLINE=97000351; PubMed=8843436;
RA      Regenbach M., Kieser H.M., Denaplatte D., Eichner A., Cullum J.,
RA      Kinashi H., Hopwood D.A.;
RT      "A set of ordered cosmids and a detailed genetic and physical map for
RT      the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL      Mol. Microbiol. 21:77-96(1996).

```

```

CC      -1- SIMILARITY: BELONGS TO THE ARAC/XYLIS FAMILY OF TRANSCRIPTIONAL
CC      REGULATORS.
DR      EMBL; AL356592; CAB92194.1; -.
DR      InterPro: IPR000005; -.
DR      Pfam; PF00165; HTH_ARAC. 1.
DR      PRINTS; PR00032; HTHARAC.
DR      PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
DR      SMART; SM00342; HTH_ARAC; 1.
KW      DNA-binding; Transcription regulation.
SQ      SEQUENCE 289 AA; 31511 MW; F709F9C93C73F56 CRC64;

Query Match      66.7%; Score 6; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3  TPVAPT 8
        |||||
Db      37  TPVAPT 42

RESULT  9
ID      09RW05      PRELIMINARY;      PRT;      293 AA.
AC      09RW05;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE      HYPOTHEICAL 30.8 KDA PROTEIN.
GN      DR0864.
OS      Deinococcus radiodurans.
OC      Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX      NCBI_TaxID=1299;
RN      [1]
RP      SEQUENCE FROM N.A..
RC      STRAIN-RI;
RA      MEDLINE=20036896; PubMed=10567266;
RA      White O., Eisen J.A., Heidelberg J.F., Hickey E.R., Peterson J.D.,
RA      Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA      Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA      Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA      Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA      Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA      Fraser C.M.;
RT      "Genome sequence of the radioresistant bacterium Deinococcus
RT      radiodurans RI."
RL      Science 286:1571-1577(1999).
DR      EMBL; AE001940; AAF10447.1; -.
DR      TIGR; DR0864; -.
KW      Hypothetical protein.
SQ      SEQUENCE 293 AA; 30849 MW; 4FA7BFEFF12A0E83 CRC64;

Query Match      66.7%; Score 6; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2  STPVAP 7
        |||||
Db      227  STPVAP 232

RESULT  10
ID      086673      PRELIMINARY;      PRT;      447 AA.
AC      086673;
DT      01-NOV-1998 (TREMBLrel. 08, Created)
DT      01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT      01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE      PUTATIVE LIPOPROTEIN.
GN      SCA42.17C.
OS      Streptomyces coelicolor.
OC      Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae;

```



OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 -RC STRAIN-A3(2);  
 RA Oliver K., Harris D.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Parthill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA MEDLINE-97000351; PubMed-8843436;  
 RX Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb streptomyces coelicolor A3(2) chromosome.";  
 RL MOL. Microbiol. 21:77-96(1996).  
 DR EMBL: AL031182; CAA20169.1; -;  
 KM Lipoprotein.  
 SQ SEQUENCE 447 AA; 46712 MW; 809E0091B7834D80 CRC64;

Query Match 66.7%; Score 6; DB 2; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPT 8  
 |||||  
 DB 63 TPVAPT 68

RESULT 11  
 O9XTQ9 PRELIMINARY; PRT; 460 AA.  
 AC O9XTQ9;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE H13N06.2 PROTEIN.  
 GN H13N06.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Leonard N.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94150718; PubMed-7906398;  
 RA Wilson R., Ahnscough R., Anderson K., Baynes C., Berke M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kersey J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons N., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,  
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL: 299942; CAB17068.1; -;  
 DR InterPro: IPR002035; -;  
 DR Pfam: PF00092; vva; 1;  
 SQ SEQUENCE 460 AA; 49872 MW; 54FA0FAC48B8AD8A CRC64;

Query Match 66.7%; Score 6; DB 5; Length 460;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPT 8  
 |||||  
 DB 73 TPVAPT 78

RESULT 12  
 ID 008484 PRELIMINARY; PRT; 637 AA.  
 AC 008484; 000028;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CHROMOSOME XV READING FRAME ORF YOR070C.  
 GN YOR29-21 OR GYPL.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bohn C., Bolotin-Fukuhara M., Daignan-Fornier B., Dang D.V.,  
 RA Valens M.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97279235; PubMed-9133743;  
 RA Valens M., Bohn C., Daignan-Fornier B., Dang V., Bolotin-Fukuhara M.;  
 RT "The sequence of a 54.7 kb fragment of yeast chromosome XV reveals the  
 presence of two tRNAs and 24 new open reading frames.";  
 RL Yeast 13:379-390(1997).  
 DR EMBL: 274978; CAA99263.1; -;  
 DR EMBL: 270678; CAA94555.1; -;  
 DR SGD: S0005596; GYPL.  
 DR InterPro: IPR00195; -;  
 DR Pfam: PF00566; TBC; 1;  
 DR SMART; SM00164; TBC; 1;  
 SQ SEQUENCE 637 AA; 73289 MW; 08C1A26B6ED37E13 CRC64;

Query Match 66.7%; Score 6; DB 3; Length 637;  
 Best Local Similarity 100.0%; Pred. No. 44;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPT 8  
 |||||  
 DB 181 TPVAPT 186

RESULT 13  
 ID 09VW35 PRELIMINARY; PRT; 652 AA.  
 AC 09VW35;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE CG8743 PROTEIN.  
 GN CG8743.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;  
RA MEDLINE-20196006; PubMed-10731132;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blake J.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Bens P.V., Berman B.P., Bhandari D., Bolintineanu S.,  
RA Borkov D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
RA Butts K.C., Butam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hosteln D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Klumel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Naylor D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacלב J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinett K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smalt T.,  
RA Spler E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RA "The genome sequence of *Drosophila melanogaster*.";  
RA Science 287:2185-2195(2000).  
DR EMBL: AE003516; AAF49118.1; -  
DR FLYBase; FBgn0036904; CG8743.  
DR InterPro; IPR001682; -  
DR InterPro; IPR002111; -  
SQ SEQUENCE 652 AA; 74251 MW; 21989FCE490235E4 CRC64;

Query Match 66.7%; Score 6; DB 5; Length 652;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STVPAP 7  
| | | | |  
Db 51 STVPAP 56

RESULT 14  
O9P5T7 PRELIMINARY; PRT; 748 AA.  
ID O9P5T7;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE GLUCAN 1, 4-ALPHA-GLUCOSIDASE RELATED PROTEIN.  
GN B5022.150.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_Taxid=5141;  
OX [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Aligh V., Hohnsels J., Brandt P., Fartmann B., Holland R.,  
RA Nykatura G., Kewes H.W., Mannhaupt G.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [12]  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL355932; CAB91434.1; -  
SQ SEQUENCE 748 AA; 79293 MW; 39316E2D4C37848A CRC64;

Query Match 66.7%; Score 6; DB 3; Length 748;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVA 6  
| | | | |  
Db 506 VSTPVA 511

RESULT 15  
O22907 PRELIMINARY; PRT; 748 AA.  
ID O22907;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE RNA HELICASE ISOLOG.  
GN T08113.17.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_Taxid=3702;  
OX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C.,  
RA Springs T.A., Mason T.M., Kerlavage A.R., Adams M.D., Somerville C.R.,  
RA Venter J.C.;  
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO DEAD/DEAH BOX HELICASE FAMILY.  
CC -1- SIMILARITY: TO HELICASE C-TERMINAL DOMAIN.  
CC -1- SIMILARITY: TO OTHER "DEAD" BOX FAMILY HELICASES.  
DR EMBL; AC002337; AAB63833.1; -  
DR Mendel; 26537; Atacth;26537.  
DR InterPro; IPR000629; -  
DR InterPro; IPR001410; -  
DR InterPro; IPR001650; -  
DR Pfam; PF002270; DEAD\_1;  
DR Pfam; PF002271; helicase\_C\_1;  
DR PROSITE; PS00039; DEAD\_ATP\_HELICASE; 1.  
DR SMART; SM00490; HELIC\_C; 1.  
KW ATP-binding; Helicase; RNA-binding.  
SQ SEQUENCE 748 AA; 81590 MW; 3F2FD0D5722D7364 CRC64;

Query Match 66.7%; Score 6; DB 10; Length 748;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPQ 9  
| | | | |  
Db 725 PVAPQ 730

RESULT 16  
O9HFZ4 PRELIMINARY; PRT; 750 AA.  
ID O9HFZ4;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE REPRessed BY TUP1 PROTEIN 1.  
GN RBT1.

OS *Candida albicans* (yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetiales; mitosporic Saccharomycetales; *Candida*.  
 OX NCBI\_TaxID=5476;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20253093; PubMed=10790384;  
 RA Braun B.R., Johnson A.D.;  
 RT "Tup1, Cph1 and Erg1 make independent contributions to filamentation  
 in *Candida albicans*.";  
 RL Genetics 155:57-67(2000).  
 RP [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20435724; PubMed=10978273;  
 RA Braun B.R., Head W.S., Wang M.X., Johnson A.D.;  
 RT "Identification and characterization of Tup1-regulated genes in  
*Candida albicans*.";  
 RL Genetics 156:31-44(2000).  
 DR EMBL: AF254142; AAG09787.1;  
 SQ SEQUENCE 750 AA; 76351 MW; 3B1C0C8EEAFCD72 CRC64;

Query Match 66.7%; Score 6; DB 3; Length 750;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STRPVAP 7  
 Db 650 STRPVAP 655

RESULT 17  
 ID 09VB49 PRELIMINARY; PRT; 775 AA.  
 AC 09VB49;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
 DE CG3361 PROTEIN.  
 GN CG3361.  
 OS *Drosophila melanogaster* (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Neilson C.R., Miklos G.L.G.,  
 RA Abill J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandal D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegyan C.,  
 RA Jalaal M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,  
 RA Lasio P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wattarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zayer J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003760; AAF5694.1;  
 DR Flybase: FBgn0039507; CG3361.  
 SQ SEQUENCE 775 AA; 85777 MW; AA3D5C6D6F871160 CRC64;

Query Match 66.7%; Score 6; DB 5; Length 775;  
 Best Local Similarity 100.0%; Pred. No. 51;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STRPVAP 7  
 Db 358 STRPVAP 363

RESULT 18  
 ID 09UAG1 PRELIMINARY; PRT; 792 AA.  
 AC 09UAG1;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE SALX-5.  
 DE Ephydaria fluviatilis.  
 OS Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;  
 OC Haptosclerida; Spongillidae; Ephydaria.  
 OX NCBI\_TaxID=31350;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99332084; PubMed=10405173;  
 RA Suga H., Ono K., Miyata T.;  
 RT "Multiple TGF-beta receptor related genes in sponge and ancient gene  
 duplication before the parazoan-eumetazoan split.";  
 RL FEBS Lett. 453:346-350(1999).  
 CC -1- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: AB026828; BAA82605.1;  
 DR InterPro: IPR000279;  
 DR InterPro: IPR002290;  
 DR Pfam: PF00069; pkinase; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 KW ATP-binding, Serine/threonine-protein kinase, Transferase.  
 SQ SEQUENCE 792 AA; 85470 MW; CEDICFOCFCE523 CRC64;

Query Match 66.7%; Score 6; DB 5; Length 792;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAPT 8  
 Db 575 TPVAPT 580

RESULT 19  
 ID 09V48 PRELIMINARY; PRT; 819 AA.  
 AC 09V48;

DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CG16807 PROTEIN.  
 GN CG16807.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 DE EPHYDROIDEA; DROSOPHILIDAE; DROSOPHILA.  
 NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaniides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paigleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 CC -1- SIMILARITY: CONTAINS A RING (C3HC4-CLASS) ZINC FINGER.  
 DR EMBL: AEO03527; AAF49474.1; -;  
 DR FlyBase: FBgn0036521; CG16807.  
 DR InterPro: IPR000571; -;  
 DR InterPro: IPR001841; -;  
 DR Pfam: PF00097; zf-C3HC4; 1.  
 DR Pfam: PF00642; zf-CCCH; 1.  
 DR PROSITE: PS00518; ZINC\_FINGER\_C3HC4; 1.  
 DR SMART: SM00184; RING; 1.  
 KW ZINC\_FINGER.  
 SQ SEQUENCE 819 AA; 90376 MW; 33DA29076EDE7E2E CRC64;

Query Match 66.7%; Score 6; DB 5; Length 819;  
 Best Local Similarity 100.0%; Pred. NO. 53;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 PVAPTO 9  
 I I I I I I  
 DB 563 PVAPTO 568

RESULT 20  
 ID Q9VEX7 PRELIMINARY; PRT; 904 AA.  
 AC Q9VEX7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CG6889 PROTEIN.  
 GN CG6889.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 DE EPHYDROIDEA; DROSOPHILIDAE; DROSOPHILA.  
 NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaniides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paigleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AEO03712; AAF55288.1; -;  
 DR FlyBase: FBgn0040071; tara.  
 DR InterPro: IPR00104; -;  
 DR PRINTS: PR00308; ANTIREEZE1.  
 SQ SEQUENCE 904 AA; 94862 MW; 084F49645F5A6998 CRC64;

Query Match 66.7%; Score 6; DB 5; Length 904;  
 Best Local Similarity 100.0%; Pred. NO. 58;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 TPVAPT 8  
 I I I I I I  
 DB 706 TPVAPT 711

## RESULT 21

Q9NHC1 PRELIMINARY; PRT; 912 AA.  
AC Q9NHC1: 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE TARA 1B ISOFORM.  
GN TARA OR CG6889.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Calgano S.T., Boube M., Cribbs D.L., Bourbon H.M.;  
RT "taranis, a novel trithorax-group of homeotic gene activators, encodes  
RT two protein isoforms related to the human cell-cycle regulator  
RT p34(Sei-1).";  
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF227212; AAF43018.1; -  
DR Flybase: FBgn0040071; tara.  
DR InterPro: IPR000104; -  
DR PRINTS: PR00308; ANTIFREEZE1.  
SQ SEQUENCE 912 AA; 95728 MW; 0DBA6C2FEE6B8F3E CRC64;

Query Match 66.7%; Score 6; DB 5; Length 912;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 TPVAPT 8  
Db 714 TPVAPT 719

## RESULT 22

Q9NHB9 PRELIMINARY; PRT; 912 AA.  
AC Q9NHB9: 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE TARA 1B ISOFORM.  
GN TARA OR CG6889.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Calgano S.T., Boube M., Cribbs D.L., Bourbon H.M.;  
RT "taranis, a novel member of the trithorax-group of homeotic gene  
RT activators from Drosophila, encodes two protein isoforms related to  
RT the human cell-cycle regulator p34(Sei-1).";  
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF227213; AAF43020.1; -  
DR Flybase: FBgn0040071; tara.  
DR InterPro: IPR000104; -  
DR PRINTS: PR00308; ANTIFREEZE1.  
SQ SEQUENCE 912 AA; 95698 MW; 1CAB7D3EFF7F544E CRC64;

Query Match 66.7%; Score 6; DB 5; Length 912;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 TPVAPT 8  
Db 714 TPVAPT 719

## RESULT 23

Q9NHC2 PRELIMINARY; PRT; 916 AA.  
AC Q9NHC2: 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE TARA 1A ISOFORM.  
GN TARA OR CG6889.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Calgano S.T., Boube M., Cribbs D.L., Bourbon H.M.;  
RT "taranis, a novel member of the trithorax-group of homeotic gene  
RT activators, encodes two protein isoforms related to the human cell-  
RT cycle regulator p34(Sei-1).";  
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF227211; AAF43017.1; -  
DR Flybase: FBgn0040071; tara.  
DR InterPro: IPR000104; -  
DR PRINTS: PR00308; ANTIFREEZE1.  
SQ SEQUENCE 916 AA; 96272 MW; A67D331E012FA39E CRC64;

Query Match 66.7%; Score 6; DB 5; Length 916;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 TPVAPT 8  
Db 718 TPVAPT 723

## RESULT 24

Q9NHC0 PRELIMINARY; PRT; 916 AA.  
AC Q9NHC0: 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE TARA 1A ISOFORM.  
GN TARA OR CG6889.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Calgano S.T., Boube M., Cribbs D.L., Bourbon H.M.;  
RT "taranis, a novel member of the trithorax-group of homeotic gene  
RT activators from Drosophila, encodes two protein isoforms related to  
RT the human cell-cycle regulator p34(Sei-1).";  
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF227213; AAF43019.1; -  
DR Flybase: FBgn0040071; tara.  
DR InterPro: IPR000104; -  
DR PRINTS: PR00308; ANTIFREEZE1.  
SQ SEQUENCE 916 AA; 96242 MW; B76C220F103B98EE CRC64;

Query Match 66.7%; Score 6; DB 5; Length 916;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 TPVAPT 8  
Db 718 TPVAPT 723

RESULT 25  
 ID 09VS50 PRELIMINARY: PRT: 1052 AA.  
 AC 09VS50  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)  
 DE CG8598 PROTEIN.  
 GN CG8598  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_Taxid=7227;  
 [1]  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 Amentides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 Suton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Abril J.F., Abmayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 Ballew K.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 Buia K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 de la Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,  
 Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 Gooder A., Gong F., Gottell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 Houston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 Jaiswal B.E., Kalish F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kohler C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mates J., McIntosh T.C., McLeod M.P., McPherson D.,  
 Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,  
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gidbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003559; AAF50579.1; -;  
 DR FlyBase: FBgn0035766; CG8598.  
 DR InterPro: IPR001969; -;  
 DR PROSITE: PS00141; ASP\_PROTEASE; UNKNOWN\_1.  
 SQ SEQUENCE 1052 AA; 117413 MW; 5DAAE81BB3B12580 CRC64;

Query Match 66.7%; Score 6; DB 5; Length 1052;  
 Best Local Similarity 100.0%; Pred. No. 65;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 STVPAP 7  
 |||||  
 Db 383 STVPAP 388

RESULT 26  
 ID 000203 PRELIMINARY: PRT: 1094 AA.  
 AC 000203  
 DT 01-JUL-1997 (TReMBLrel. 04, Created)  
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)  
 DE BETA-ADAPTIN 3A (AP-3 COMPLEX BETA3A SUBUNIT) (CEREBELLAR DEGENERATION  
 ANTIEN BETA-NAP HOMOLOG BETA-3A).  
 GN ADTB3A.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_Taxid=9606;  
 [1]  
 RN  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=HEART;  
 RX MEDLINE=97296491; PubMed=9151686;  
 RA Simpson F., Peden A.A., Christopoulou L., Robinson M.S.;  
 RT "Characterization of the adaptor-related protein complex, AP-3";  
 RL J. Cell Biol. 137:835-845(1997).  
 [2]  
 RN  
 RP SEQUENCE FROM N.A., PHOSPHORYLATION, AND TISSUE SPECIFICITY.  
 RC TISSUE=PANCREAS;  
 RX MEDLINE=97326075; PubMed=9182526;  
 RA Dell'Angelica E.C., Ooi C.E., Bonifacio J.S.;  
 RT "Beta3A-adaptin, a subunit of the adaptor-like complex AP-3";  
 RL J. Biol. Chem. 272:15078-15084(1997).  
 [3]  
 RN  
 RP HERMANSKY-PUDLAK SYNDROME (HPS) DISEASE.  
 RC TISSUE=SKIN FIBROBLAST, AND PERIPHERAL BLOOD;  
 RX MEDLINE=99149014; PubMed=10024875;  
 RA Dell'Angelica E.C., Shetler V., Aguilar R.C., Gahl W.A.,  
 Bonifacio J.S.;  
 RT "Altered trafficking of lysosomal proteins in Hermansky-Pudlak  
 syndrome due to mutations in the beta 3A subunit of the AP-3  
 adaptor";  
 RL Mol. Cell 3:11-21(1999).  
 [4]  
 RN  
 RP FUNCTION: PART OF THE AP-3 COMPLEX, AN ADAPTOR-RELATED COMPLEX  
 WHICH IS NOT CLATHRIN-ASSOCIATED. THE COMPLEX IS ASSOCIATED WITH  
 THE GOLGI REGION AS WELL AS MORE PERIPHERAL STRUCTURES. IT  
 FACILITATES THE BUDDING OF VESICLES FROM THE GOLGI MEMBRANE AND  
 MAY BE DIRECTLY INVOLVED IN TRAFFICKING TO LYOSOMES.  
 CC -1 SUBUNIT: ASSEMBLY PROTEIN COMPLEX 3 (AP-3) IS A HETEROTETRAMER  
 COMPOSED OF TWO LARGE CHAINS (DELTA AND BETA3), A MEDIUM CHAIN  
 (WU3) AND A SMALL CHAIN (SIGMA3). EXPRESSED.  
 CC -1 TISSUE SPECIFICITY: UBQUITOUSLY EXPRESSED.  
 CC -1 PTM: PHOSPHORYLATED ON SERINE.  
 CC -1 DISEASE: DEFECTS IN ADTB3A ARE THE CAUSE OF HERMANSKY-PUDLAK  
 SYNDROME (HPS), AN AUTOSOMAL RECESSIVE DISORDER CHARACTERIZED BY  
 OCULOCUTANEOUS ALBINISM AND PLATELET STORAGE POOL DEFICIENCY.  
 CC -1 SIMILARITY: BELONGS TO THE ADAPTER COMPLEXES LARGE SUBUNITS  
 FAMILY.  
 CC EMBL: U01931; AAD03778.1; -;  
 DR EMBL: U01504; AAB61638.1; -;  
 DR MIM: 203300; -;  
 DR MIM: 603401; -;  
 DR InterPro: IPR002553; -;  
 DR Pfam: PF01602; Adaptin\_N; 1.  
 KW Albinism; Disease mutation; Golgi stack; phosphorylation;  
 KW Polymorphism; Protein transport; transport.  
 FT VARIANT 390 410 MISSING (IN HPS).  
 FT VARIANT 580 580 L -> R (IN HPS).  
 FT CONFLICT 804 804 MISSING (IN REF. 1).  
 SQ SEQUENCE 1094 AA; 121349 MW; AC683CE18FE3EB6 CRC64;

Query Match 66.7%; Score 6; DB 4; Length 1094;  
 Best Local Similarity 100.0%; Pred. No. 67;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VSTPVA 6

Db 827 VSTPVA 832

|||||

```
RESULT 27
ID Q921T1 PRELIMINARY; PRT: 1105 AA.
AC Q921T1:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE AP-3 COMPLEX BETA3A SUBUNIT (ADAPTOR-RELATED PROTEIN COMPLEX AP-3 BETA
DE 1 SUBUNIT).
GN AP3b1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/HEJ;
RA Feng L., Seymour A.B., Jiang S.Y., To A., Peden A.A., Novak E.K.,
RA Zhen L., Rusniak M.E., Elcher E.M., Robinson M.S., Gorin M.B.,
RA Swank R.T.;
RT "The Beta3A Subunit Gene (Ap3b1) of the AP-3 Complex is Altered in the
RT Mouse Hypopigmentation Mutant Pearl, a Model for Hermansky Pudlak
RT Syndrome and Night Blindness.";
RL Hum. Mol. Genet. 0:0-0(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129S6/SVEVYAC;
RX MEDLINE=20512043; PubMed=11056055;
RA Feng L., Rigatelli B.W., Novak E.K., Gorin M.B., Swank R.T.;
RT "Genomic Structure of the Mouse Ap3b1 Gene in Normal and Pearl Mice.";
RL Genomics 69:370-379(2000).
DR EMBL; AF103809; AAC78338.1; -.
DR EMBL; AF255589; AAG23622.1; JOINED.
DR EMBL; AF255586; AAG23622.1; JOINED.
DR EMBL; AF255567; AAG23622.1; JOINED.
DR EMBL; AF255568; AAG23622.1; JOINED.
DR EMBL; AF255569; AAG23622.1; JOINED.
DR EMBL; AF255570; AAG23622.1; JOINED.
DR EMBL; AF255571; AAG23622.1; JOINED.
DR EMBL; AF255572; AAG23622.1; JOINED.
DR EMBL; AF255573; AAG23622.1; JOINED.
DR EMBL; AF255574; AAG23622.1; JOINED.
DR EMBL; AF255575; AAG23622.1; JOINED.
DR EMBL; AF255576; AAG23622.1; JOINED.
DR EMBL; AF255577; AAG23622.1; JOINED.
DR EMBL; AF255578; AAG23622.1; JOINED.
DR EMBL; AF255579; AAG23622.1; JOINED.
DR EMBL; AF255580; AAG23622.1; JOINED.
DR EMBL; AF255581; AAG23622.1; JOINED.
DR EMBL; AF255582; AAG23622.1; JOINED.
DR EMBL; AF255583; AAG23622.1; JOINED.
DR EMBL; AF255584; AAG23622.1; JOINED.
DR EMBL; AF255585; AAG23622.1; JOINED.
DR EMBL; AF255586; AAG23622.1; JOINED.
DR EMBL; AF255587; AAG23622.1; JOINED.
DR EMBL; AF255588; AAG23622.1; JOINED.
DR MGD; MGI:133879; Ap3b1.
DR InterPro; IPR000130; -.
DR InterPro; IPR002553; -.
DR Pfam; PF01602; Adaplin_N; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1105 AA; 122869 MW; 586B818CE4FB5ABE CRC64;
```

Query Match 66.7%; Score 6; DB 11; Length 1105;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVA 6

Db 838 VSTPVA 843

|||||

```
RESULT 28
ID Q9LBG3 PRELIMINARY; PRT: 1653 AA.
AC Q9LBG3:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE PAA.
GN PAA.
OS Streptococcus criceti.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1333;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E49;
RA Tamura H., Kato H.;
RT "Cell surface antigen I/II - Streptococcus cricetus.";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB042239; BAA95000.1; -.
DR InterPro; IPR001899; -.
DR InterPro; IPR002965; -.
DR Pfam; PF00746; Gram_Pos_anchor; 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
SQ SEQUENCE 1653 AA; 180862 MW; 93F28E215BF598DE CRC64;
```

Query Match 66.7%; Score 6; DB 2; Length 1653;  
Best Local Similarity 100.0%; Pred. No. 94;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPQ 9  
Db 955 PVAPQ 960

```
RESULT 29
ID Q9PVZ2 PRELIMINARY; PRT: 2037 AA.
AC Q9PVZ2:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE NUCLEOPORIN CAN.
GN CAN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99384295; PubMed=10454574;
RA Askjaer P., Bachl A., Wilm M., Bischoff R., Weeks D.L., Ogniewski V.,
RA Ohno M., Niehrs C., Kjems J., Mattei I.W., Forrester M.;
RT "Rangp-Regulated Interactions of CRM1 with Nucleoporins and a
RT shuttling DEAD-box helicase.";
RL Mol. Cell. Biol. 19:6276-6285(1999).
DR EMBL; AJ243889; CAB53357.1; -.
DR InterPro; IPR000515; -.
DR InterPro; IPR001680; -.
DR PROSITE; PS00402; BPD_TRANSP_INN_MEMBER; UNKNOWN_1.
DR SMART; SM00320; WD40; 1.
DR Pfam; PF01602; Adaplin_N; 1.
SQ SEQUENCE 2037 AA; 208931 MW; 225DEA9A0435635 CRC64;
```

Query Match 66.7%; Score 6; DB 13; Length 2037;

Best Local Similarity 100.0%; Pred. NO. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVAP 7  
|||||  
DB 1052 STPVAP 1057

RESULT 30

O9YIV2 PRELIMINARY; PRT; 49 AA.

ID O9YIV2; 09YIV2;

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE PEDIBIN PRECURSOR.

GN PER.

OS Hydra magnipapillata (Hydra), and

OS Hydra attenuata (Hydra) (Hydra vulgaris).

OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroids; Anthomedusae;

OC Hydridae; Hydra.

OX NCBI\_Taxid=6085, 6087;

RN [1]

RP SEQUENCE FROM N.A.

RC SPECIES-H.magnipapillata; STRAIN-105;

RA Takahashi T., Hotta M., Fujisawa T.;

RT "Foot dominant expression of the foot activating peptide, Hym-346/Pedibin, in Hydra.";

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC SPECIES-H.magnipapillata; STRAIN-105;

RA MEDLINE=99094877; PubMed=9876180;

RA Greens A., Shimizu H., Hoffmeister S., Bode H., Fujisawa T.;

RT "The novel signal peptides, pedibin and Hym-346, lower positional

RT value thereby enhancing foot formation in hydra.";

RL Development 126:517-524(1999).

RN [3]

RP SEQUENCE FROM N.A.

RC SPECIES-H.attenuata;

RX MEDLINE=96232307; PubMed=8674432;

RA Hoffmeister S.A.H.;

RT "Isolation and characterization of two new morphogenetically active

RT peptides from Hydra vulgaris.";

RL Development 122:1941-1948(1996).

RN [4]

RP SEQUENCE FROM N.A.

RC SPECIES-H.attenuata;

RA Hoffmeister S.A.H., Herrmann D.;

RT "Cloning of the foot-formation activating peptide pedibin from Hydra

RT vulgaris.";

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE FROM N.A.

RC SPECIES-H.attenuata; S.A.H., Herrmann D.;

RA Hoffmeister-Ullrich S.A.H., Herrmann D.;

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB030084; BAA82554.1; -;

DR EMBL; AF177908; AAF01282.1; -;

KM Signal.

FT CHAIN 1 28

FT SIGNAL 29 49

FT CHAIN 29 49

FT SIGNAL 29 49

FT CHAIN 29 49

FT SIGNAL 29 49

FT CHAIN 29 49

Query Match 55.6%; Score 5; DB 5; Length 49;

Best Local Similarity 100.0%; Pred. No. 69;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 STPV 5  
|||||  
DB 7 STPV 11

RESULT 31

O47006

ID O47006; PRELIMINARY; PRT; 62 AA.

AC O47006;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)

DE FLAGELLIN (FRAGMENT).

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_Taxid=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=045:H23 (ECRC H23);

RA Kwang J., Wilson R., Yang S., He T.;

RL Clin. Diagn. Lab. Immunol. 0:0-0(0).

DR EMBL; U57313; AAB01993.1; -;

FT NON\_TER 1 1

FT NON\_TER 62 62

FT NON\_TER 62 62

SO SEQUENCE 62 AA; 6217 MW; 6BF545C9F2134414 CRC64;

OY 3 TPVAP 7  
|||||  
DB 15 TPVAP 19

RESULT 32

O9RRH1

ID O9RRH1; PRELIMINARY; PRT; 75 AA.

AC O9RRH1;

DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE HYPOTHETICAL 7.7 KDA PROTEIN.

GN DR2520.

OS Deinococcus radiodurans.

OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.

OX NCBI\_Taxid=1299;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RI;

RX MEDLINE=20036896; PubMed=10567266;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.;

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.;

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.;

RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.;

RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.;

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.;

RA Fraser C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus

RT radiodurans RI.";

RL Science 286:1571-1577(1999).

DR EMBL; AE002081; AAF12064.1; -;

DR TIGR; DR2520; -;

KM Hypothetical protein.

FT SIGNAL 75 AA; 7704 MW; EBE47B7BAE33B0E CRC64;

OY 5 VAPTO 9  
|||||  
DB 24 VAPTO 28

Query Match 55.6%; Score 5; DB 2; Length 75;

Best Local Similarity 100.0%; Pred. No. 98;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

RESULT 33
ID 017728 PRELIMINARY; PRT: 82 AA.
AC 017728;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE C06E4.2 PROTEIN.
GN C06E4.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaidon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mleg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX Du Z., Gattung S.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX Waterston R.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U41277; AAA82477.1;
SQ SEQUENCE 82 AA; 9058 MW; E6C6AC722EBB125B6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 5; Length 82;
Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 VAPTO 9
DB 40 VAPTO 44

RESULT 34
ID P91335 PRELIMINARY; PRT: 88 AA.
AC P91335;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE C05MID F55C7.
GN C05MID F55C7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaidon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mleg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX Du Z., Gattung S.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX Waterston R.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U41277; AAA82477.1;
SQ SEQUENCE 82 AA; 9058 MW; E6C6AC722EBB125B6 CRC64;

```

```

RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaidon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierly-Mleg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RA Du Z., Le T.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U80436; AAC71107.1;
SQ SEQUENCE 88 AA; 9771 MW; 59FB61FDD7C0CE3 CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 5; Length 88;
Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6
DB 31 STPVA 35

RESULT 35
ID 067673 PRELIMINARY; PRT: 94 AA.
AC 067673;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE HEXON PROTEIN (FRAGMENT).
OS porcine adenovirus sp.
OS Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=47901;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LUNG;
RA Pfling-Akerdion P.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: X95630; CAA64884.1;
DR HSSP: P03277; IDHX.
DR InterPro: IPR000736;
DR Pfam: PF01065; Adeno_hexon; 1.
KW Hexon Protein.
FT NON_TER 1 94
FT NON_TER 94 94
SQ SEQUENCE 94 AA; 10807 MW; 2646FBE97DDCC9FD CRC64;

Query Match
Best Local Similarity 100.0%; Score 5; DB 14; Length 94;
Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 VAPTO 9
DB 39 VAPTO 43

RESULT 36
ID 090YD6 PRELIMINARY; PRT: 95 AA.
AC 090YD6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

```

DE SMALL INDUCIBLE CYTOKINE A12 PRECURSOR.  
GN SCYAL2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=STL/J; TISSUE=SPLEEN;  
RX MEDLINE=99370037; PubMed=10438970;  
RA Teuscher C., Butterfield R.J., Ma R.Z., Zachary J.F., Doerge R.W.,  
Blankenhorn E.P.;  
RT "Sequence polymorphisms in the chemokines Scy1 (TCA-3), Scy2  
RT (monocyte chemoattractant protein (MCP)-1), and Scyal2 (MCP-5) are  
RT candidates for ea27, a locus controlling susceptibility to monophasic  
RT remitting/nonrelapsing experimental allergic encephalomyelitis.";  
RL J. Immunol. 163:2262-2266(1999).  
DR EMBL; AF065938; AAF15388.1; -.  
DR HSSP; P13500; IDOL.  
DR MGD; MGI:108224; Scyal2.  
DR InterPro: IPR000476; -.  
DR InterPro: IPR000827; -.  
DR InterPro: IPR001811; -.  
DR Pfam; PF00048; IL8; 1.  
DR ProDom; PD002047; -. 1.  
DR PROSITE; PS00472; SMALL\_CYTOKINES\_CC; 1.  
DR SMART; SM00199; SCY; 1.  
SQ SEQUENCE 95 AA; 10727 MW; 9DB56B029729D0B0 CRC64;

Query Match 55.6%; Score 5; DB 11; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTPV 5  
DB 27 VSTPV 31

RESULT 37  
068302 PRELIMINARY; PRT; 95 AA.  
AC 068302;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E1 (GP32) (GP35)]  
DE (FRAGMENT).  
GN E1.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HCV-BB43;  
RA Songsiyilai S., Kanistanon D.;  
RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
DR EMBL; U23396; AAA64862.1; -.  
DR InterPro: IPR002519; -.  
DR Pfam; PF01539; HCV\_env; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
KW Transmembrane.  
FT NON\_TER 1  
SQ SEQUENCE 95 AA; 10305 MW; 86270AA69397533A CRC64;

Query Match 55.6%; Score 5; DB 14; Length 95;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PVAPT 8  
DB 23 PVAPT 27

RESULT 38  
068301 PRELIMINARY; PRT; 102 AA.  
AC 068301;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E1 (GP32) (GP35)]  
DE (FRAGMENT).  
GN E1.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HCV-BB8;  
RA Songsiyilai S., Kanistanon D.;  
RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.  
DR EMBL; U23395; AAA64861.1; -.  
DR InterPro: IPR002519; -.  
DR Pfam; PF01539; HCV\_env; 1.  
KW Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
KW Transmembrane.  
FT NON\_TER 1  
SQ SEQUENCE 102 AA; 11040 MW; 2C560825E0AD043E CRC64;

Query Match 55.6%; Score 5; DB 14; Length 102;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PVAPT 8  
DB 24 PVAPT 28

RESULT 39  
093516 PRELIMINARY; PRT; 108 AA.  
AC 093516;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE SALMOSIN2 (FRAGMENT).  
OS Gloydius halys brevicaudus.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;  
OC Viperidae; Crotalinae; Gloydius.  
OX NCBI\_TaxID=66175;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=VENOM GLAND;  
RA Park D.S., Kim H.D., Chung K.H., Kim D.S., Yun Y.D.;  
RT "Cloning and Characterization of Novel Disintegrins from Agkistrodon  
RT halys venom.";  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF055337; AAC42597.1; -.  
DR HSSP; P18619; IFVL.  
DR InterPro: IPR001762; -.  
DR Pfam; PF00200; disintegrin; 1.  
DR PRINTS; PR00289; DISINTEGRIN.  
DR PROSITE; PS50214; DISINTEGRIN\_2; 1.  
DR SMART; SM00050; DISIN; 1.  
FT NON\_TER 1  
SQ SEQUENCE 108 AA; 11726 MW; 573F0B6918970C8C CRC64;

Query Match  
Best Local Similarity 55.6%; Score 5; DB 13; Length 108;  
Matches 5; Conservative 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
11111  
DB 25 VSTPV 29

## RESULT 40

097949 PRELIMINARY; PRT; 114 AA.  
AC 097949:  
DT 01-MAY-1999 (TREMBLREL. 10, Created)  
DT 01-MAY-1999 (TREMBLREL. 10, Last sequence update)  
DE BETA-MICROSEMINOPROTEIN PRECURSOR.  
GN MSP-J1.  
OS Saguinus oedipus (Cotton-top tamarin).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.  
OX NCBI\_TaxID=9490;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lundwall A., Makinen M.;  
RT "Molecular cloning of genes encoding beta-microseminoprotein from the  
RT cotton-top tamarin."  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ010155; CAB38123.1; -.  
DR EMBL; AJ010155; CAB38123.1; JOINED.  
DR EMBL; AJ010157; CAB38123.1; JOINED.  
KW Signal.  
FT SIGNAL.  
SQ SEQUENCE 114 AA; 12746 MW; F3F05B013445BAD4 CRC64;

Query Match  
Best Local Similarity 55.6%; Score 5; DB 6; Length 114;  
Matches 5; Conservative 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
11111  
DB 73 VSTPV 77

## RESULT 41

097935 PRELIMINARY; PRT; 114 AA.  
AC 097935:  
DT 01-MAY-1999 (TREMBLREL. 10, Created)  
DT 01-MAY-1999 (TREMBLREL. 10, Last sequence update)  
DE BETA-MICROSEMINOPROTEIN PRECURSOR.  
GN MSP-E1.  
OS Saguinus oedipus (Cotton-top tamarin).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.  
OX NCBI\_TaxID=9490;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lundwall A., Makinen M.;  
RT "Molecular cloning of genes encoding beta-microseminoprotein from the  
RT cotton-top tamarin."  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ010154; CAB38105.1; -.  
KW Signal.  
FT SIGNAL.  
SQ SEQUENCE 114 AA; 12738 MW; 05E7A410125C94B2 CRC64;

Query Match  
Best Local Similarity 55.6%; Score 5; DB 6; Length 114;  
Matches 5; Conservative 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
11111  
DB 73 VSTPV 77

## RESULT 42

056724 PRELIMINARY; PRT; 115 AA.  
AC 056724; 056723;  
DT 01-NOV-1996 (TREMBLREL. 01, Created)  
DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)  
DE DELTA-SUBUNIT, METHYLMALONYL-COA DECARBOXYLASE.  
OS Veillonella parvula.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Sporomusa subbranch;  
OC Veillonella.  
OX NCBI\_TaxID=29466;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Huder J.B., Dimroth P.;  
RL J. Biol. Chem. 0:0-0(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-94043308; PubMed-8227015;  
RA Huder J.B., Dimroth P.;  
RT "Sequence of the sodium ion pump methylmalonyl-CoA decarboxylase from  
RT Veillonella parvula."  
RL J. Biol. Chem. 268:24564-24571(1993).  
DR EMBL; Z24754; CA80873.1; -.  
DR EMBL; U22208; AAC36821.1; -.  
SQ SEQUENCE 115 AA; 11951 MW; 3F873586E596EC26 CRC64;

Query Match  
Best Local Similarity 55.6%; Score 5; DB 2; Length 115;  
Matches 5; Conservative 100.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
11111  
DB 60 TPVAP 64

## RESULT 43

09W2V1 PRELIMINARY; PRT; 115 AA.  
AC 09W2V1:  
DT 01-MAY-2000 (TREMBLREL. 13, Created)  
DT 01-MAY-2000 (TREMBLREL. 13, Last sequence update)  
DE 01-JUN-2000 (TREMBLREL. 14, Last annotation update)  
DE CG12640 PROTEIN.  
GN CG12640.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE-20196006; PubMed-10731132;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Chaney M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abail J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Folsler C., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harlin N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iobagwan C.,  
 RA Jallat B., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nussken D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spler E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrstkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenhach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gldbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003450; AAF6587.1;  
 DR FlyBase: FBgn0030202; CG12640.  
 DR InterPro: IPR001745;  
 DR PROSITE: PS00627; GHMP\_KINASES\_ATP; UNKNOWN\_1.  
 SQ SEQUENCE 115 AA; 12840 MW; 7975280AB7AFBD79 CRC64;

Query Match 55.6%; Score 5; DB 5; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
 |||||  
 DB 32 TPVAP 36

RESULT 44  
 090222 PRELIMINARY: PRT: 115 AA.

AC 090222;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE PREPRO-HALYSTATIN 3 (FRAGMENT).  
 OS Agkistrodon halys pallas (Chinese water moccasin).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroides;  
 OC Viperidae; Crotalinae; Gloydius.  
 OC NCBI\_Taxid=8714;  
 OX [1]  
 RN  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER.  
 RA Fujisawa Y., Kuroda S., Notoya K., Konishi H., Terashita Z.,  
 RT "Halystatin, a novel disintegrin from agkistrodon halys, is a potent  
 inhibitor of bone resorption and platelet aggregation.";  
 RL Takeeda Kenkyusho Ho 53:39-56(1994).  
 CC -1- STIMILARITY: COMPAINS A DISINTEGRIN DOMAIN.  
 DR EMBL: D28871; BAA06026.1;  
 DR HSSP: P18619; IFLV.  
 DR MEROPS: M12.134;  
 DR InterPro: IPR001590;  
 DR InterPro: IPR001762;  
 DR Pfam: PF00200; disintegrin; 1.  
 DR PRINTS: PR00289; DISINTEGRIN.

DR PROSITE: PS50215; ADAM\_MEROP; 1.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; 1.  
 DR PROSITE: PS50214; DISINTEGRIN\_2; 1.  
 DR SMART: SM00050; DISIN: 1.  
 KW Blood coagulation; Platelet; Cell adhesion; Venom.  
 FT NON\_TER 1  
 SQ SEQUENCE 115 AA; 12389 MW; 00CD85D94635A274 CRC64;

Query Match 55.6%; Score 5; DB 13; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 |||||  
 DB 34 VSTPV 38

RESULT 45  
 09X0K8 PRELIMINARY: PRT: 117 AA.

DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE CONSERVED HYPOTHETICAL PROTEIN.  
 GN TM1124.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogales; Thermotoga.  
 OX NCBI\_Taxid=2336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MSB8 / DSM 3109;  
 RX MEDLINE=99287316; PubMed=10360571;  
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,  
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
 genome sequence of *Thermotoga maritima*."  
 RL Nature 399:323-329(1999).  
 DR EMBL: AE001770; AAD36200.1;  
 DR TIGR: TM1124;  
 SQ SEQUENCE 117 AA; 13692 MW; BDB44AF940E1A95 CRC64;

Query Match 55.6%; Score 5; DB 2; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
 |||||  
 DB 30 STPVA 34

RESULT 46  
 09NZE0 PRELIMINARY: PRT: 117 AA.

AC 09NZE0;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
 DE BM-016.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BONE MARROW;  
 RA Zhao M., Song H., Li N., Peng Y., Han Z., Chen Z.;

RT "A novel gene expressed in human bone marrow."  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF208858; AAF64272.1; -  
 SQ SEQUENCE 117 AA: 12834 MW: D7A64D8378BD7983 CRC64:

Query Match 55.6%; Score 5; DB 4; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 |||||  
 DB 43 VSTPV 47

RESULT 47  
 O9VWMO PRELIMINARY; PRT; 117 AA.  
 AC O9VWMO:  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE CG15883 PROTEIN.  
 GN CG15883.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Adair J.F., Abdayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhargava D., Bolshakov S.,  
 RA Borovica D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Modary C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Releert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Slier B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spirdling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter L., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Weissman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195 (2000).  
 DR EMBL: AE003511; AAF48918.1; -  
 DR Flybase; FBgn0030985; CG15883.

SQ SEQUENCE 117 AA: 12897 MW: 085AA0062082827E CRC64:

Query Match 55.6%; Score 5; DB 5; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
 |||||  
 DB 35 STPVA 39

RESULT 48  
 O90221 PRELIMINARY; PRT; 117 AA.  
 AC O90221:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE PREPRO-HALYSTATTIN 2 (FRAGMENT).  
 OS Agkistrodon halys palas (Chinese water moccasin).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidae;  
 OC Viperae; Crotalinae; Gloydius.  
 OX NCBI\_TaxID=8714;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Fujisawa Y., Kuroda S., Notoya K., Konishi H., Terashita Z.,  
 RT "Halystatin, a novel disintegrin from agkistrodon halys, is a potent  
 RT inhibitor of bone resorption and platelet aggregation.";  
 RL Takeda Kenkyusho Ho 53:39-56(1994).  
 CC -1- SIMILARITY: CONTRAINS A DISINTEGRIN DOMAIN.  
 DR EMBL: D28871; BA06027.1; -  
 DR HSP: P18619; IFLV.  
 DR MEROPS: M12.134; -  
 DR Interpro: IPR001590; -  
 DR Interpro: IPR001762; -  
 DR Pfam: PF00200; disintegrin.1.  
 DR PRINTS: PR00289; DISINTEGRIN.  
 DR PROSITE: PS00215; ADAM\_MEROP; 1.  
 DR PROSITE: PS00427; DISINTEGRIN.1; 1.  
 DR PROSITE: PS00214; DISINTEGRIN.2; 1.  
 DR SMART: SM00050; DISIN.1.  
 KW Blood coagulation; Platelet; Cell adhesion; Venom.  
 FT NON\_TER 1  
 SQ SEQUENCE 117 AA: 12686 MW: B1P9A98056BAB07B CRC64:

Query Match 55.6%; Score 5; DB 13; Length 117;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
 |||||  
 DB 34 VSTPV 38

RESULT 49  
 O9N3V1 PRELIMINARY; PRT; 121 AA.  
 AC O9N3V1:  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE HYPOTHETICAL PROTEIN Y47D7A.H.  
 GN Y47D7A.H.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode *C. elegans*: a platform for  
 investigating biology. The *C. elegans* Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 DR EMBL: AC024790; AAF0636.1; -  
 SQ SEQUENCE 121 AA; 13441 MW; FF21FD197434C60F CRC64;

Query Match 55.6%; Score 5; DB 5; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
 |||||  
 Db 46 PVAPT 50

RESULT 50  
 O23980 PRELIMINARY; PRT; 121 AA.  
 AC O23980;  
 DT 01-JAN-1998 (TREMBLER1.05, Created)  
 DT 01-JAN-1998 (TREMBLER1.05, Last sequence update)  
 DT 01-NOV-1999 (TREMBLER1.12, Last annotation update)  
 DE COLD-REGULATED PROTEIN (FRAGMENT).  
 OS Hordeum vulgare (Barley).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae;  
 OC Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. GIORGIE; TISSUE-LEAF;  
 RA Cattiveili L.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. GIORGIE; TISSUE-LEAF;  
 RA Cattiveili L.; Bartels D.;  
 RL Plant Physiol. 93:1504-1510(1990).  
 DR EMBL: AJ000100; CAA03925.1; -  
 DR Mendel; 25606; Horvu;1290;25606.  
 FT NON\_TER 1  
 SQ SEQUENCE 121 AA; 12484 MW; 1B2A44C69C189735 CRC64;

Query Match 55.6%; Score 5; DB 10; Length 121;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
 |||||  
 Db 84 TPVAP 88

Search completed: August 15, 2001, 12:42:32  
 Job time: 517 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2001, 12:32:55 ; Search time 31.32 seconds  
(without alignments)  
5.917 Million cell updates/sec

Title: US-09-372-036-26

Perfect score: 9

Sequence: 1 VSTPVPAPQ 9

Scoring table:

OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	100.0	9	2	US-08-456-670B-26
2	9	100.0	21	1	US-08-127-499A-27
3	9	100.0	21	1	US-08-482-847-27
4	9	100.0	232	2	US-08-456-670B-39
5	9	100.0	478	2	US-08-456-670B-40
6	9	100.0	484	1	US-08-127-499A-26
7	9	100.0	484	1	US-08-482-847-26
8	8	88.9	12	2	US-08-456-670B-30
9	6	66.7	10	2	US-08-456-670B-31
10	6	66.7	20	2	US-08-456-670B-17
11	5	55.6	11	4	US-09-001-984C-55
12	5	55.6	94	1	US-07-899-535A-1
13	5	55.6	192	1	US-08-086-428B-57
14	5	55.6	192	1	US-08-440-103-44
15	5	55.6	192	1	US-08-440-542-44
16	5	55.6	192	1	US-08-231-368-44
17	5	55.6	192	1	US-08-440-210-44
18	5	55.6	192	2	US-08-468-570-57
19	5	55.6	192	2	US-08-290-665A-57
20	5	55.6	192	4	US-08-444-818-167
21	5	55.6	192	5	PCR-US95-10398-57
22	5	55.6	264	3	US-08-969-644-12
23	5	55.6	264	3	US-08-444-189-12
24	5	55.6	264	4	US-08-465-465-2
25	5	55.6	264	4	US-08-468-544-12
26	5	55.6	286	1	US-08-382-184-3
27	5	55.6	286	2	US-08-641-356-3

28	5	55.6	286	4	US-09-132-528-4	Sequence 4, Appl
29	5	55.6	286	4	US-08-875-494-3	Sequence 3, Appl
30	5	55.6	325	1	US-08-382-184-2	Sequence 2, Appl
31	5	55.6	325	2	US-08-641-356-2	Sequence 2, Appl
32	5	55.6	325	4	US-09-132-528-2	Sequence 2, Appl
33	5	55.6	325	4	US-09-132-528-3	Sequence 3, Appl
34	5	55.6	325	4	US-08-875-494-2	Sequence 2, Appl
35	5	55.6	339	4	US-08-444-818-152	Sequence 152, App
36	5	55.6	339	4	US-08-444-818-156	Sequence 156, App
37	5	55.6	359	1	US-08-137-627-4	Sequence 4, Appl
38	5	55.6	359	2	US-08-865-348-4	Sequence 4, Appl
39	5	55.6	370	3	US-08-969-644-23	Sequence 23, Appl
40	5	55.6	370	3	US-08-444-189-23	Sequence 23, Appl
41	5	55.6	370	4	US-08-468-544-23	Sequence 23, Appl
42	5	55.6	377	2	US-08-853-659A-41	Sequence 41, Appl
43	5	55.6	384	3	US-08-852-824-4	Sequence 4, Appl
44	5	55.6	396	1	US-08-769-309A-15	Sequence 15, Appl
45	5	55.6	396	3	US-08-994-570-15	Sequence 15, Appl
46	5	55.6	401	2	US-08-839-008-5	Sequence 5, Appl
47	5	55.6	422	1	US-08-403-545-4	Sequence 4, Appl
48	5	55.6	422	4	US-08-404-381-4	Sequence 4, Appl
49	5	55.6	428	1	US-08-353-550-1	Sequence 1, Appl
50	5	55.6	428	2	US-08-551-687-1	Sequence 1, Appl
51	5	55.6	431	3	US-08-807-342B-5	Sequence 5, Appl
52	5	55.6	434	2	US-08-710-249-4	Sequence 4, Appl
53	5	55.6	453	1	US-08-769-309A-14	Sequence 14, Appl
54	5	55.6	453	3	US-08-994-570-14	Sequence 14, Appl
55	5	55.6	468	2	US-08-839-008-7	Sequence 7, Appl
56	5	55.6	468	4	US-09-032-523-8	Sequence 8, Appl
57	5	55.6	469	1	US-08-353-550-6	Sequence 6, Appl
58	5	55.6	469	2	US-08-551-687-6	Sequence 6, Appl
59	5	55.6	486	3	US-08-889-841-8	Sequence 8, Appl
60	5	55.6	491	3	US-08-889-841-10	Sequence 10, Appl
61	5	55.6	498	1	US-08-496-885A-6	Sequence 6, Appl
62	5	55.6	498	2	US-08-466-589-12	Sequence 12, Appl
63	5	55.6	498	2	US-08-700-636-11	Sequence 12, Appl
64	5	55.6	498	3	US-08-467-574-12	Sequence 12, Appl
65	5	55.6	501	2	US-08-845-566-1	Sequence 1, Appl
66	5	55.6	501	3	US-08-991-408-4	Sequence 4, Appl
67	5	55.6	605	2	US-08-687-936A-1	Sequence 1, Appl
68	5	55.6	651	1	US-08-769-309A-17	Sequence 17, Appl
69	5	55.6	651	3	US-08-994-570-17	Sequence 17, Appl
70	5	55.6	655	1	US-07-736-178C-2	Sequence 2, Appl
71	5	55.6	670	2	US-08-366-547-2	Sequence 2, Appl
72	5	55.6	671	2	US-08-737-716-13	Sequence 13, Appl
73	5	55.6	685	3	US-08-872-855-2	Sequence 2, Appl
74	5	55.6	739	3	US-09-035-648-24	Sequence 24, Appl
75	5	55.6	739	4	US-09-001-925-24	Sequence 24, Appl
76	5	55.6	788	1	US-08-572-225-1	Sequence 1, Appl
77	5	55.6	933	4	US-08-764-870-14	Sequence 14, Appl
78	5	55.6	1009	2	US-08-680-326-31	Sequence 31, Appl
79	5	55.6	1013	2	US-08-866-650-5	Sequence 5, Appl
80	5	55.6	1013	3	US-09-021-287-5	Sequence 5, Appl
81	5	55.6	1013	2	US-08-991-408-2	Sequence 2, Appl
82	5	55.6	1148	2	US-08-313-185-58	Sequence 58, Appl
83	5	55.6	1148	2	US-09-082-614A-58	Sequence 58, Appl
84	5	55.6	1213	1	US-08-188-582-20	Sequence 20, Appl
85	5	55.6	1213	1	US-08-646-715-20	Sequence 20, Appl
86	5	55.6	1222	2	US-08-682-517-15	Sequence 15, Appl
87	5	55.6	1253	2	US-08-682-517-9	Sequence 9, Appl
88	5	55.6	1253	1	US-07-920-281C-3	Sequence 3, Appl
89	5	55.6	1253	4	US-08-466-277-3	Sequence 3, Appl
90	5	55.6	1449	3	US-08-840-062-6	Sequence 6, Appl
91	5	55.6	1466	2	US-08-687-956A-23	Sequence 23, Appl
92	5	55.6	1627	1	US-07-665-792E-9	Sequence 9, Appl
93	5	55.6	1780	1	US-08-769-309A-5	Sequence 5, Appl
94	5	55.6	1780	3	US-08-994-570-5	Sequence 5, Appl
95	5	55.6	2476	2	US-08-276-967-2	Sequence 2, Appl
96	5	55.6	3418	2	US-08-639-501-2	Sequence 2, Appl
97	5	55.6	3418	2	US-08-603-753D-4	Sequence 4, Appl
98	5	55.6	3418	3	US-09-044-946-2	Sequence 2, Appl
99	5	55.6	3418	3	US-08-755-587-44	Sequence 44, Appl
100	5	55.6	3418	3	US-09-044-908-2	Sequence 2, Appl

101	5	55.6	3418	4	US-09-099-753-4	Sequence 4, Appl1	174	4	44.4	28	2	US-08-392-816-4	Sequence 4, Appl1
102	5	55.6	3418	4	US-08-986-106-4	Sequence 4, Appl1	175	4	44.4	31	1	US-07-829-462-4	Sequence 4, Appl1
103	4	44.4	4	4	US-08-973-462-28	Sequence 28, Appl1	176	4	44.4	31	1	US-08-340-812-4	Sequence 4, Appl1
104	4	44.4	7	1	US-08-261-206A-18	Sequence 18, Appl1	177	4	44.4	31	1	US-08-459-064B-4	Sequence 4, Appl1
105	4	44.4	7	2	US-08-350-260A-421	Sequence 421, App	178	4	44.4	31	1	US-08-488-252-14	Sequence 34, Appl1
106	4	44.4	8	2	US-08-769-745-11	Sequence 11, Appl1	179	4	44.4	31	2	US-08-460-421A-4	Sequence 4, Appl1
107	4	44.4	8	3	US-08-335-733D-38	Sequence 38, Appl1	180	4	44.4	31	2	US-08-023-980B-30	Sequence 30, Appl1
108	4	44.4	8	3	US-08-335-733D-39	Sequence 39, Appl1	181	4	44.4	31	2	US-08-486-953A-25	Sequence 25, Appl1
109	4	44.4	8	3	US-08-335-733D-40	Sequence 40, Appl1	182	4	44.4	31	3	US-08-466-368-16	Sequence 16, Appl1
110	4	44.4	8	3	US-08-335-733D-41	Sequence 41, Appl1	183	4	44.4	31	5	PCT-US93-00909-4	Sequence 4, Appl1
111	4	44.4	8	3	US-08-335-733D-42	Sequence 42, Appl1	184	4	44.4	31	6	5512648-2	Patent No. 5512648
112	4	44.4	9	1	US-08-338-634-7	Sequence 7, Appl1	185	4	44.4	33	1	US-08-086-428B-137	Sequence 137, App
113	4	44.4	9	1	US-08-244-855-14	Sequence 14, Appl1	186	4	44.4	33	2	US-08-468-570-137	Sequence 137, App
114	4	44.4	9	3	US-08-335-733D-5	Sequence 5, Appl1	187	4	44.4	33	2	US-08-468-570-137	Sequence 137, App
115	4	44.4	9	3	US-09-001-984C-17	Sequence 17, Appl1	188	4	44.4	33	5	US-08-290-665A-241	Sequence 241, App
116	4	44.4	10	4	US-08-836-075A-188	Sequence 188, App	189	4	44.4	35	4	PCT-US95-10398-241	Sequence 241, App
117	4	44.4	11	1	US-08-336-343A-29	Sequence 29, Appl1	190	4	44.4	35	4	US-09-001-984C-11	Sequence 11, Appl1
118	4	44.4	11	1	US-08-338-634-8	Sequence 8, Appl1	191	4	44.4	37	1	US-08-001-984C-42	Sequence 42, Appl1
119	4	44.4	11	1	US-08-436-670B-35	Sequence 35, Appl1	192	4	44.4	37	1	US-08-463-660-8	Sequence 8, Appl1
120	4	44.4	11	2	US-08-456-670B-38	Sequence 38, Appl1	193	4	44.4	37	1	US-08-678-280-8	Sequence 8, Appl1
121	4	44.4	11	3	US-08-705-875A-11	Sequence 11, Appl1	194	4	44.4	38	1	US-08-176-500-85	Sequence 85, Appl1
122	4	44.4	11	3	US-08-652-877-29	Sequence 29, Appl1	195	4	44.4	38	1	US-08-471-052A-85	Sequence 85, Appl1
123	4	44.4	11	4	US-08-476-515A-29	Sequence 29, Appl1	196	4	44.4	38	1	US-08-189-331-85	Sequence 85, Appl1
124	4	44.4	11	4	US-09-001-984C-7	Sequence 7, Appl1	197	4	44.4	38	2	US-08-471-039-85	Sequence 85, Appl1
125	4	44.4	11	4	US-09-001-984C-23	Sequence 23, Appl1	198	4	44.4	38	2	US-08-471-800-85	Sequence 85, Appl1
126	4	44.4	11	5	PCT-US95-16415-18	Sequence 18, Appl1	199	4	44.4	38	2	US-08-468-161-3	Sequence 85, Appl1
127	4	44.4	12	4	US-08-602-959A-255	Sequence 255, App	200	4	44.4	38	2	US-08-146-028-18	Sequence 18, Appl1
128	4	44.4	13	3	US-08-705-875A-12	Sequence 12, Appl1	201	4	44.4	38	2	US-08-166-068-85	Sequence 85, Appl1
129	4	44.4	13	4	US-09-001-984C-8	Sequence 8, Appl1	202	4	44.4	38	3	US-09-273-685-3	Sequence 3, Appl1
130	4	44.4	13	4	US-09-001-984C-9	Sequence 9, Appl1	203	4	44.4	38	4	US-08-723-425A-18	Sequence 18, Appl1
131	4	44.4	13	4	US-09-001-984C-20	Sequence 20, Appl1	204	4	44.4	38	4	US-09-112-206-18	Sequence 18, Appl1
132	4	44.4	13	4	US-09-001-984C-25	Sequence 25, Appl1	205	4	44.4	38	5	PCT-US95-11934-3	Sequence 3, Appl1
133	4	44.4	14	3	US-08-705-875A-13	Sequence 13, Appl1	206	4	44.4	39	2	US-08-455-625-3	Sequence 3, Appl1
134	4	44.4	15	1	US-08-408-604A-31	Sequence 31, Appl1	207	4	44.4	39	2	US-08-455-625-29	Sequence 29, Appl1
135	4	44.4	15	2	US-08-687-956A-3	Sequence 3, Appl1	208	4	44.4	39	4	US-08-455-685-3	Sequence 3, Appl1
136	4	44.4	15	2	US-08-687-956A-6	Sequence 6, Appl1	209	4	44.4	39	5	US-08-455-685-29	Sequence 29, Appl1
137	4	44.4	15	2	US-08-602-959A-410	Sequence 410, App	210	4	44.4	39	5	PCT-US94-05142-29	Sequence 29, Appl1
138	4	44.4	15	5	PCT-US93-09626-31	Sequence 31, Appl1	211	4	44.4	42	1	US-08-450-945-61	Sequence 61, Appl1
139	4	44.4	16	1	US-08-447-925-5	Sequence 5, Appl1	212	4	44.4	42	1	US-08-976-161-61	Sequence 61, Appl1
140	4	44.4	16	2	US-08-102-385G-35	Sequence 35, Appl1	213	4	44.4	48	3	US-08-665-259-6	Sequence 6, Appl1
141	4	44.4	16	3	US-08-705-875A-14	Sequence 14, Appl1	214	4	44.4	48	3	US-08-762-500-6	Sequence 6, Appl1
142	4	44.4	16	4	US-08-602-999A-179	Sequence 179, App	215	4	44.4	49	3	US-08-665-259-5	Sequence 5, Appl1
143	4	44.4	18	1	US-08-441-534A-6	Sequence 6, Appl1	216	4	44.4	49	3	US-08-762-500-5	Sequence 5, Appl1
144	4	44.4	18	1	US-08-629-752-6	Sequence 6, Appl1	217	4	44.4	52	1	US-08-294-189-17	Sequence 17, Appl1
145	4	44.4	18	2	US-08-802-991-6	Sequence 6, Appl1	218	4	44.4	53	2	US-08-238-821B-53	Sequence 53, Appl1
146	4	44.4	19	2	US-08-284-391B-39	Sequence 39, Appl1	219	4	44.4	53	2	PCT-US95-05744-53	Sequence 53, Appl1
147	4	44.4	19	2	US-08-284-391B-41	Sequence 41, Appl1	220	4	44.4	54	2	US-08-400-159-16	Sequence 16, Appl1
148	4	44.4	20	1	US-08-218-025A-72	Sequence 72, Appl1	221	4	44.4	54	2	US-08-611-729A-16	Sequence 16, Appl1
149	4	44.4	20	2	US-08-934-915-10	Sequence 10, Appl1	222	4	44.4	58	2	US-08-152-721B-19	Sequence 19, Appl1
150	4	44.4	21	1	US-08-273-776-1	Sequence 1, Appl1	223	4	44.4	59	1	US-08-306-871-25	Sequence 25, Appl1
151	4	44.4	21	5	PCT-US92-10432-1	Sequence 1, Appl1	224	4	44.4	59	1	US-08-569-959-25	Sequence 25, Appl1
152	4	44.4	22	2	US-08-833-546-9	Sequence 9, Appl1	225	4	44.4	59	3	US-08-651-136C-66	Sequence 66, Appl1
153	4	44.4	22	4	US-09-388-664-9	Sequence 9, Appl1	226	4	44.4	60	1	US-08-447-925-1	Sequence 1, Appl1
154	4	44.4	23	1	US-08-218-025A-132	Sequence 132, App	227	4	44.4	60	1	US-08-099-354-7	Sequence 7, Appl1
155	4	44.4	23	2	US-08-493-235-40	Sequence 40, Appl1	228	4	44.4	60	2	US-08-288-059-3	Sequence 3, Appl1
156	4	44.4	23	2	US-08-492-027A-3	Sequence 3, Appl1	229	4	44.4	61	1	US-07-734-534A-5	Sequence 5, Appl1
157	4	44.4	23	3	US-08-485-324-15	Sequence 15, Appl1	230	4	44.4	65	4	US-09-188-930-297	Sequence 297, App
158	4	44.4	23	3	US-08-447-506-15	Sequence 15, Appl1	231	4	44.4	69	1	US-08-137-800-40	Sequence 40, Appl1
159	4	44.4	23	3	US-08-235-437-15	Sequence 15, Appl1	232	4	44.4	69	1	US-08-477-383-40	Sequence 40, Appl1
160	4	44.4	23	3	US-08-447-515-15	Sequence 15, Appl1	233	4	44.4	69	1	US-08-487-174-40	Sequence 40, Appl1
161	4	44.4	24	1	US-08-488-252-31	Sequence 31, Appl1	234	4	44.4	69	1	US-08-487-174-40	Sequence 40, Appl1
162	4	44.4	24	2	US-08-407-252-4	Sequence 4, Appl1	235	4	44.4	70	2	US-08-935-450-12	Sequence 12, Appl1
163	4	44.4	24	3	US-08-705-875A-15	Sequence 15, Appl1	236	4	44.4	80	1	US-08-137-800-41	Sequence 41, Appl1
164	4	44.4	25	1	US-08-451-405A-3	Sequence 3, Appl1	237	4	44.4	80	1	US-08-477-383-11	Sequence 11, Appl1
165	4	44.4	25	2	US-08-474-696A-4	Sequence 4, Appl1	238	4	44.4	80	1	US-08-477-174-41	Sequence 41, Appl1
166	4	44.4	25	2	US-08-455-625-25	Sequence 25, Appl1	239	4	44.4	80	1	US-08-480-750-41	Sequence 41, Appl1
167	4	44.4	25	4	US-08-455-685-25	Sequence 25, Appl1	240	4	44.4	80	1	US-08-341-219-23	Sequence 23, Appl1
168	4	44.4	25	5	PCT-US94-05142-25	Sequence 25, Appl1	241	4	44.4	80	4	US-08-912-114A-23	Sequence 23, Appl1
169	4	44.4	26	1	US-08-383-753-26	Sequence 26, Appl1	242	4	44.4	82	1	US-08-225-757B-12	Sequence 12, Appl1
170	4	44.4	26	2	US-08-586-772-26	Sequence 26, Appl1	243	4	44.4	82	1	US-08-446-038B-14	Sequence 14, Appl1
171	4	44.4	26	2	US-08-959-512-26	Sequence 26, Appl1	244	4	44.4	82	1	US-08-446-010B-14	Sequence 14, Appl1
172	4	44.4	26	3	US-08-705-875A-16	Sequence 16, Appl1	245	4	44.4	82	2	US-08-805-445-14	Sequence 14, Appl1
173	4	44.4	26	3	US-08-335-733D-1	Sequence 1, Appl1	246	4	44.4	82	2	US-08-064-067D-14	Sequence 14, Appl1



247	4	44.4	82	2	US-09-066-208-14	Sequence 14, Appl	320	4	44.4	124	5	PCR-US95-08743-131	Sequence 131, App
248	4	44.4	83	1	US-08-370-225-21	Sequence 21, Appl	321	4	44.4	127	2	US-08-137-717D-37	Sequence 37, Appl
249	4	44.4	83	1	US-08-461-859-21	Sequence 21, Appl	322	4	44.4	127	2	US-08-436-717-37	Sequence 37, Appl
250	4	44.4	83	5	PCR-US93-10069-21	Sequence 21, Appl	323	4	44.4	138	1	US-08-686-878A-14	Sequence 14, Appl
251	4	44.4	84	3	US-08-648-342-10	Sequence 10, Appl	324	4	44.4	139	1	US-07-718-274A-2	Sequence 2, Appl
252	4	44.4	94	1	US-08-167-035-23	Sequence 23, Appl	325	4	44.4	139	1	US-08-163-877-10	Sequence 10, Appl
253	4	44.4	94	1	US-08-208-887A-23	Sequence 23, Appl	326	4	44.4	139	1	US-08-149-106-2	Sequence 2, Appl
254	4	44.4	94	2	US-08-539-005-23	Sequence 23, Appl	327	4	44.4	139	1	US-08-298-021-2	Sequence 2, Appl
255	4	44.4	94	2	US-08-238-821B-46	Sequence 46, Appl	328	4	44.4	139	1	US-08-278-729A-5	Sequence 5, Appl
256	4	44.4	94	5	PCR-US95-05744-46	Sequence 46, Appl	329	4	44.4	139	1	US-08-278-729A-6	Sequence 6, Appl
257	4	44.4	100	3	US-09-034-916-13	Sequence 13, Appl	330	4	44.4	139	1	US-08-155-343A-5	Sequence 5, Appl
258	4	44.4	100	4	US-08-973-462-11	Sequence 11, Appl	331	4	44.4	139	1	US-08-155-343A-6	Sequence 6, Appl
259	4	44.4	102	1	US-08-335-583C-47	Sequence 47, Appl	332	4	44.4	139	1	US-08-406-672-5	Sequence 5, Appl
260	4	44.4	102	2	US-08-808-982-8	Sequence 8, Appl	333	4	44.4	139	1	US-08-406-672-6	Sequence 6, Appl
261	4	44.4	102	2	US-08-288-508C-18	Sequence 18, Appl	334	4	44.4	139	1	US-08-643-563A-5	Sequence 5, Appl
262	4	44.4	102	3	US-08-478-097A-16	Sequence 16, Appl	335	4	44.4	139	1	US-08-643-563A-6	Sequence 6, Appl
263	4	44.4	102	3	US-08-289-222E-27	Sequence 27, Appl	336	4	44.4	139	1	US-08-643-763A-5	Sequence 5, Appl
264	4	44.4	102	4	US-09-054-526B-27	Sequence 27, Appl	337	4	44.4	139	1	US-08-643-763A-6	Sequence 6, Appl
265	4	44.4	102	4	US-08-931-858E-161	Sequence 161, App	338	4	44.4	139	1	US-08-462-623-5	Sequence 5, Appl
266	4	44.4	102	4	US-08-981-739-161	Sequence 161, App	339	4	44.4	139	1	US-08-462-623-6	Sequence 6, Appl
267	4	44.4	104	1	US-07-764-731B-8	Sequence 8, Appl	340	4	44.4	139	1	US-08-451-953A-5	Sequence 5, Appl
268	4	44.4	105	1	US-08-422-101-9	Sequence 9, Appl	341	4	44.4	139	1	US-08-451-953A-6	Sequence 6, Appl
269	4	44.4	105	1	US-08-422-091-9	Sequence 9, Appl	342	4	44.4	139	1	US-08-360-914B-10	Sequence 10, Appl
270	4	44.4	105	2	US-08-422-092-9	Sequence 9, Appl	343	4	44.4	139	1	US-08-741-589A-10	Sequence 10, Appl
271	4	44.4	105	2	US-08-788-800-6	Sequence 6, Appl	344	4	44.4	139	2	US-08-445-468A-5	Sequence 5, Appl
272	4	44.4	105	3	US-08-422-093-9	Sequence 9, Appl	345	4	44.4	139	2	US-08-445-468A-6	Sequence 6, Appl
273	4	44.4	105	3	US-08-422-112-9	Sequence 9, Appl	346	4	44.4	139	2	US-08-461-357A-5	Sequence 5, Appl
274	4	44.4	107	2	US-08-810-572A-4	Sequence 4, Appl	347	4	44.4	139	2	US-08-461-357A-6	Sequence 6, Appl
275	4	44.4	109	1	US-07-829-462-1	Sequence 1, Appl	348	4	44.4	139	2	US-08-912-088-5	Sequence 5, Appl
276	4	44.4	109	1	US-08-340-812-1	Sequence 1, Appl	349	4	44.4	139	2	US-08-912-088-6	Sequence 6, Appl
277	4	44.4	109	1	US-08-459-064B-1	Sequence 1, Appl	350	4	44.4	139	3	US-08-278-730A-5	Sequence 5, Appl
278	4	44.4	109	2	US-08-460-421A-1	Sequence 1, Appl	351	4	44.4	139	3	US-08-278-730A-6	Sequence 6, Appl
279	4	44.4	109	2	US-08-717-169-1	Sequence 1, Appl	352	4	44.4	139	3	US-08-445-467-5	Sequence 5, Appl
280	4	44.4	109	2	US-08-761-227A-51	Sequence 51, Appl	353	4	44.4	139	3	US-08-445-467-6	Sequence 6, Appl
281	4	44.4	109	3	US-08-646-322-5	Sequence 5, Appl	354	4	44.4	139	3	US-08-443-676-1	Sequence 1, Appl
282	4	44.4	109	5	PCR-US93-00909-1	Sequence 1, Appl	355	4	44.4	139	3	US-08-480-515A-5	Sequence 5, Appl
283	4	44.4	110	1	US-07-849-389-7	Sequence 7, Appl	356	4	44.4	139	3	US-08-480-515A-6	Sequence 6, Appl
284	4	44.4	111	3	US-08-545-809A-111	Sequence 111, App	357	4	44.4	139	4	US-08-414-033A-5	Sequence 5, Appl
285	4	44.4	113	4	US-08-836-075A-70	Sequence 70, Appl	358	4	44.4	139	4	US-08-414-033A-6	Sequence 6, Appl
286	4	44.4	113	4	US-08-836-075A-72	Sequence 72, Appl	359	4	44.4	139	4	US-08-271-556A-3	Sequence 3, Appl
287	4	44.4	113	4	US-08-836-075A-74	Sequence 74, Appl	360	4	44.4	139	4	US-08-271-556A-4	Sequence 4, Appl
288	4	44.4	113	4	US-08-836-075A-78	Sequence 78, Appl	361	4	44.4	139	4	US-08-905-223-44	Sequence 44, App
289	4	44.4	114	4	US-09-188-930-161	Sequence 161, App	362	4	44.4	139	5	PCR-US92-01968-5	Sequence 5, Appl
290	4	44.4	114	4	US-09-188-930-288	Sequence 288, App	363	4	44.4	139	5	PCR-US92-01968-6	Sequence 6, Appl
291	4	44.4	115	2	US-07-903-029-6	Sequence 6, Appl	364	4	44.4	139	5	PCR-US93-07190-5	Sequence 5, Appl
292	4	44.4	115	3	US-08-478-097A-30	Sequence 30, Appl	365	4	44.4	139	5	PCR-US93-07190-6	Sequence 6, Appl
293	4	44.4	116	1	US-08-636-253-2	Sequence 2, Appl	366	4	44.4	139	5	PCR-US93-07231-5	Sequence 5, Appl
294	4	44.4	117	1	US-08-448-196A-2	Sequence 2, Appl	367	4	44.4	139	5	PCR-US93-08742-5	Sequence 5, Appl
295	4	44.4	118	1	US-08-481-377-12	Sequence 12, Appl	368	4	44.4	139	5	PCR-US93-08742-6	Sequence 6, Appl
296	4	44.4	118	2	US-08-491-835-10	Sequence 10, Appl	369	4	44.4	139	5	PCR-US93-08742-6	Sequence 6, Appl
297	4	44.4	118	3	US-08-392-794A-4	Sequence 4, Appl	370	4	44.4	139	5	PCR-US93-08808-5	Sequence 5, Appl
298	4	44.4	118	3	US-09-153-733A-12	Sequence 12, Appl	371	4	44.4	139	5	PCR-US93-08808-6	Sequence 6, Appl
299	4	44.4	118	3	US-08-946-092A-10	Sequence 10, Appl	372	4	44.4	139	5	PCR-US93-08885-5	Sequence 5, Appl
300	4	44.4	118	4	US-09-172-062-10	Sequence 10, Appl	373	4	44.4	139	5	PCR-US93-08885-6	Sequence 6, Appl
301	4	44.4	118	5	PCR-US94-00665-12	Sequence 12, Appl	374	4	44.4	139	5	PCR-US94-13181-10	Sequence 10, Appl
302	4	44.4	118	5	PCR-US94-00665-10	Sequence 10, Appl	375	4	44.4	142	2	US-08-805-117-3	Sequence 3, Appl
303	4	44.4	119	1	US-08-581-552B-12	Sequence 12, Appl	376	4	44.4	151	2	US-08-722-050-8	Sequence 8, Appl
304	4	44.4	119	1	US-08-455-559-18	Sequence 18, Appl	377	4	44.4	153	3	US-08-851-843A-177	Sequence 177, App
305	4	44.4	119	2	US-08-525-596B-22	Sequence 22, Appl	378	4	44.4	153	4	US-08-974-549A-296	Sequence 296, App
306	4	44.4	119	2	US-08-581-528A-12	Sequence 12, Appl	379	4	44.4	154	2	US-08-330-394A-29	Sequence 29, Appl
307	4	44.4	119	3	US-09-097-616-12	Sequence 12, Appl	380	4	44.4	155	3	US-08-950-720A-10	Sequence 10, Appl
308	4	44.4	119	3	US-09-177-860A-22	Sequence 22, Appl	381	4	44.4	156	2	US-08-330-394A-22	Sequence 22, Appl
309	4	44.4	119	4	US-09-156-316-9	Sequence 9, Appl	382	4	44.4	158	1	US-08-611-107-4	Sequence 4, Appl
310	4	44.4	119	4	US-08-624-635-13	Sequence 13, Appl	383	4	44.4	158	2	US-08-422-560A-4	Sequence 4, Appl
311	4	44.4	119	4	US-09-145-060-18	Sequence 18, Appl	384	4	44.4	158	4	US-08-468-793-4	Sequence 4, Appl
312	4	44.4	119	5	PCR-US94-00657-18	Sequence 18, Appl	385	4	44.4	161	1	US-08-450-257-56	Sequence 56, Appl
313	4	44.4	119	5	PCR-US94-07762-12	Sequence 12, Appl	386	4	44.4	161	1	US-08-450-246-56	Sequence 56, Appl
314	4	44.4	119	5	PCR-US94-07799-12	Sequence 12, Appl	387	4	44.4	161	1	US-08-450-098-56	Sequence 56, Appl
315	4	44.4	121	6	530466-3	Sequence 1, Appl	388	4	44.4	161	1	US-08-451-223-56	Sequence 56, Appl
316	4	44.4	122	6	530466-3	Sequence 1, Appl	389	4	44.4	161	1	US-08-450-236-56	Sequence 56, Appl
317	4	44.4	124	1	US-08-276-852-131	Sequence 131, App	390	4	44.4	161	2	US-08-621-803-249	Sequence 249, App
318	4	44.4	124	1	US-08-899-575-131	Sequence 131, App	391	4	44.4	161	3	US-09-045-764A-4	Sequence 4, Appl
319	4	44.4	124	1	US-08-899-575-131	Sequence 131, App	392	4	44.4	161	4	US-09-271-970-2	Sequence 2, Appl

393	4	44.4	152	1	US-08-048-164A-2	Sequence 2, App11
394	4	44.4	152	1	US-08-460-462-2	Sequence 2, App11
395	4	44.4	152	1	US-08-460-457-2	Sequence 2, App11
396	4	44.4	152	1	US-08-460-458-2	Sequence 2, App11
397	4	44.4	152	1	US-08-460-455-2	Sequence 2, App11
398	4	44.4	152	2	US-08-330-394A-2	Sequence 2, App11
399	4	44.4	152	2	US-08-983-045-2	Sequence 2, App11
400	4	44.4	153	3	US-09-006-636-7	Sequence 7, App11
401	4	44.4	153	3	US-09-006-632-7	Sequence 7, App11
402	4	44.4	156	2	US-08-483-695-3	Sequence 3, App11
403	4	44.4	156	2	US-07-965-285-3	Sequence 3, App11
404	4	44.4	156	2	US-08-487-231-3	Sequence 3, App11
405	4	44.4	156	4	US-09-201-912-3	Sequence 3, App11
406	4	44.4	159	1	US-07-841-646-9	Sequence 9, App11
407	4	44.4	159	1	US-08-147-023-9	Sequence 9, App11
408	4	44.4	159	1	US-08-447-570-9	Sequence 9, App11
409	4	44.4	159	2	US-08-449-700-9	Sequence 9, App11
410	4	44.4	159	2	US-08-449-699A-9	Sequence 9, App11
411	4	44.4	174	2	US-08-683-562B-45	Sequence 45, App1
412	4	44.4	174	2	US-08-683-562B-48	Sequence 48, App1
413	4	44.4	176	3	US-08-392-794A-8	Sequence 8, App11
414	4	44.4	176	6	5185440-7	Patent No. 5185440
415	4	44.4	177	1	US-08-450-257-61	Sequence 61, App1
416	4	44.4	177	1	US-08-450-246-61	Sequence 61, App1
417	4	44.4	177	1	US-08-450-098-61	Sequence 61, App1
418	4	44.4	177	1	US-08-451-233-61	Sequence 61, App1
419	4	44.4	177	1	US-08-450-236-61	Sequence 61, App1
420	4	44.4	178	3	US-08-705-875A-8	Sequence 8, App11
421	4	44.4	178	3	US-09-122-443-9	Sequence 9, App11
422	4	44.4	178	4	US-09-220-731-23	Sequence 23, App1
423	4	44.4	178	4	US-09-220-731-24	Sequence 24, App1
424	4	44.4	178	4	US-09-211-970-8	Sequence 8, App11
425	4	44.4	179	2	US-08-621-803-257	Sequence 257, App
426	4	44.4	179	2	US-08-621-803-261	Sequence 261, App
427	4	44.4	180	1	US-07-953-230A-12	Sequence 12, App1
428	4	44.4	180	6	5405942-4	Patent No. 5405942
429	4	44.4	181	4	US-08-961-083-114	Sequence 114, App
430	4	44.4	186	1	US-08-928-843-3	Sequence 3, App11
431	4	44.4	186	3	US-09-129-055-3	Sequence 3, App11
432	4	44.4	187	1	US-08-450-257-62	Sequence 62, App1
433	4	44.4	187	1	US-08-450-246-62	Sequence 62, App1
434	4	44.4	187	1	US-08-450-098-62	Sequence 62, App1
435	4	44.4	187	1	US-08-451-233-62	Sequence 62, App1
436	4	44.4	187	1	US-08-450-236-62	Sequence 62, App1
437	4	44.4	190	3	US-08-779-764A-18	Sequence 18, App1
438	4	44.4	190	3	US-09-271-970A-14	Sequence 14, App1
439	4	44.4	192	1	US-08-086-428B-53	Sequence 53, App1
440	4	44.4	192	1	US-08-086-428B-58	Sequence 58, App1
441	4	44.4	192	1	US-08-086-428B-85	Sequence 85, App1
442	4	44.4	192	2	US-08-468-570-53	Sequence 53, App1
443	4	44.4	192	2	US-08-468-570-58	Sequence 58, App1
444	4	44.4	192	2	US-08-468-570-85	Sequence 85, App1
445	4	44.4	192	2	US-08-290-665A-53	Sequence 53, App1
446	4	44.4	192	2	US-08-290-665A-58	Sequence 58, App1
447	4	44.4	192	2	US-08-290-665A-85	Sequence 85, App1
448	4	44.4	192	4	US-09-475-316A-23	Sequence 23, App1
449	4	44.4	192	5	PCT-US94-05795-3	Sequence 3, App11
450	4	44.4	192	5	PCT-US95-10398-53	Sequence 53, App1
451	4	44.4	192	5	PCT-US95-10398-58	Sequence 58, App1
452	4	44.4	192	5	PCT-US95-10398-85	Sequence 85, App1
453	4	44.4	193	2	US-08-536-626A-39	Sequence 39, App1
454	4	44.4	194	2	US-08-621-803-255	Sequence 255, App
455	4	44.4	195	2	US-08-621-803-263	Sequence 263, App
456	4	44.4	196	4	US-09-383-586-11	Sequence 11, App1
457	4	44.4	201	2	US-08-531-525-13	Sequence 13, App1
458	4	44.4	201	2	US-08-916-901-8	Sequence 8, App11
459	4	44.4	201	2	US-08-916-901-8	Sequence 8, App11
460	4	44.4	201	2	US-08-918-270A-13	Sequence 13, App1
461	4	44.4	208	4	US-08-097-869-6	Sequence 6, App11
462	4	44.4	210	1	US-07-667-276A-2	Sequence 2, App11
463	4	44.4	212	1	US-08-973-462-25	Sequence 25, App1
464	4	44.4	214	1	US-08-217-327-4	Sequence 4, App11
465	4	44.4	215	2	US-08-912-129A-58	Sequence 58, App1
466	4	44.4	216	2	US-08-929-418-2	Sequence 2, App11
467	4	44.4	220	1	US-08-726-306A-29	Sequence 29, App1
468	4	44.4	221	1	US-08-032-848C-13	Sequence 13, App1
469	4	44.4	221	2	US-08-861-659-6	Sequence 6, App11
470	4	44.4	221	2	US-09-134-596-6	Sequence 6, App11
471	4	44.4	221	3	US-09-293-273-6	Sequence 6, App11
472	4	44.4	221	3	US-09-293-273-6	Sequence 6, App11
473	4	44.4	224	1	US-08-553-516-2	Sequence 2, App11
474	4	44.4	225	2	US-08-886-765-2	Sequence 2, App11
475	4	44.4	225	4	US-09-115-660-2	Sequence 2, App11
476	4	44.4	229	4	US-08-751-359-22	Sequence 22, App1
477	4	44.4	233	4	US-08-523-894-6	Sequence 6, App11
478	4	44.4	234	4	US-08-487-550-2	Sequence 2, App11
479	4	44.4	235	2	US-08-378-939-12	Sequence 12, App1
480	4	44.4	235	4	US-09-049-672A-10	Sequence 10, App1
481	4	44.4	235	4	US-09-049-672A-12	Sequence 12, App1
482	4	44.4	235	4	US-09-329-350-31	Sequence 31, App1
483	4	44.4	236	3	US-08-487-550-10	Sequence 10, App1
484	4	44.4	236	4	US-09-049-672A-7	Sequence 7, App11
485	4	44.4	236	4	US-08-961-083-96	Sequence 96, App1
486	4	44.4	237	2	US-08-672-564-3	Sequence 3, App11
487	4	44.4	238	4	US-09-216-295-12	Sequence 12, App1
488	4	44.4	240	3	US-09-248-335-42	Sequence 42, App1
489	4	44.4	240	4	US-09-049-672A-11	Sequence 11, App1
490	4	44.4	248	1	US-08-338-130-2	Sequence 48, App1
491	4	44.4	248	2	US-08-921-426-4	Sequence 2, App11
492	4	44.4	248	3	US-08-816-915-4	Sequence 4, App11
493	4	44.4	248	5	PCT-US95-07743-4	Sequence 4, App11
494	4	44.4	249	1	US-08-450-257-57	Sequence 57, App1
495	4	44.4	249	1	US-08-450-246-57	Sequence 57, App1
496	4	44.4	249	1	US-08-450-098-57	Sequence 57, App1
497	4	44.4	249	1	US-08-451-233-57	Sequence 57, App1
498	4	44.4	249	1	US-08-450-236-57	Sequence 57, App1
499	4	44.4	249	2	US-08-991-946A-11	Sequence 11, App1
500	4	44.4	251	4	US-08-944-483-28	Sequence 28, App1
501	4	44.4	251	2	US-08-310-912A-106	Sequence 106, App
502	4	44.4	252	3	US-08-841-089-106	Sequence 106, App
503	4	44.4	255	5	PCT-US95-04570-106	Sequence 106, App
504	4	44.4	255	5	PCT-US95-04589-106	Sequence 106, App
505	4	44.4	257	2	US-08-715-204-6	Sequence 6, App11
506	4	44.4	257	3	US-09-162-597-6	Sequence 6, App11
507	4	44.4	257	3	US-07-721-761A-29	Sequence 29, App1
508	4	44.4	258	1	US-07-978-687-29	Sequence 29, App1
509	4	44.4	258	4	US-08-961-083-90	Sequence 90, App1
510	4	44.4	258	5	PCT-US91-05801-29	Sequence 29, App1
511	4	44.4	259	4	US-09-216-295-5	Sequence 5, App11
512	4	44.4	262	2	US-07-857-224B-64	Sequence 64, App1
513	4	44.4	264	2	US-08-494-907-6	Sequence 6, App11
514	4	44.4	264	5	PCT-US96-10986-6	Sequence 6, App11
515	4	44.4	267	4	US-08-454-928A-10	Sequence 10, App1
516	4	44.4	269	1	US-08-241-766-4	Sequence 4, App11
517	4	44.4	269	1	US-08-241-766-5	Sequence 5, App11
518	4	44.4	269	1	US-08-241-766-6	Sequence 6, App11
519	4	44.4	269	1	US-08-241-766-7	Sequence 7, App11
520	4	44.4	269	1	US-08-241-766-14	Sequence 14, App1
521	4	44.4	269	2	US-08-484-905-116	Sequence 116, App
522	4	44.4	269	3	US-08-481-985B-116	Sequence 116, App
523	4	44.4	269	3	US-08-370-476-116	Sequence 116, App
524	4	44.4	270	2	US-09-031-485-7	Sequence 7, App11
525	4	44.4	270	2	US-08-847-429A-7	Sequence 7, App11
526	4	44.4	270	2	US-07-857-224B-31	Sequence 31, App1
527	4	44.4	270	2	US-09-065-474-7	Sequence 7, App11
528	4	44.4	272	2	US-08-160-524A-7	Sequence 7, App11
529	4	44.4	272	2	US-08-492-027A-4	Sequence 4, App11
530	4	44.4	275	2	US-08-578-709-13	Sequence 13, App1
531	4	44.4	278	1	US-07-921-807B-10	Sequence 10, App1
532	4	44.4	278	1	US-08-441-944A-10	Sequence 10, App1
533	4	44.4	282	5	PCT-US95-13335-1	Sequence 1, App11
534	4	44.4	283	3	US-08-658-469-2	Sequence 2, App11
535	4	44.4	283	3	US-09-081-420-2	Sequence 2, App11
536	4	44.4	284	2	US-08-767-096-3	Sequence 3, App11
537	4	44.4	287	4	US-09-065-474-145	Sequence 145, App
538	4	44.4	293	2	US-08-810-572A-2	Sequence 2, App11

539	4	44.4	293	3	US-09-203-716-2	Sequence 2, Appli	612	4	44.4	303	3	US-08-876-398A-20	Sequence 20, Appl
540	4	44.4	294	2	US-08-874-347-26	Sequence 26, Appl	613	4	44.4	303	3	US-08-876-398A-22	Sequence 22, Appl
541	4	44.4	294	3	US-09-093-522-26	Sequence 26, Appl	614	4	44.4	303	3	US-08-876-398A-24	Sequence 24, Appl
542	4	44.4	297	1	US-08-176-620A-16	Sequence 16, Appl	615	4	44.4	303	3	US-08-876-398A-26	Sequence 26, Appl
543	4	44.4	297	2	US-08-874-347-23	Sequence 23, Appl	616	4	44.4	303	3	US-08-876-398A-28	Sequence 28, Appl
544	4	44.4	297	2	US-08-874-347-24	Sequence 24, Appl	617	4	44.4	303	3	US-08-876-398A-30	Sequence 30, Appl
545	4	44.4	297	2	US-08-461-985-16	Sequence 16, Appl	618	4	44.4	303	3	US-08-876-398A-32	Sequence 32, Appl
546	4	44.4	297	3	US-09-093-522-23	Sequence 23, Appl	619	4	44.4	303	3	US-08-876-398A-34	Sequence 34, Appl
547	4	44.4	298	3	US-09-093-522-24	Sequence 24, Appl	620	4	44.4	303	3	US-08-876-398A-36	Sequence 36, Appl
548	4	44.4	298	1	US-08-348-792-4	Sequence 4, Appli	621	4	44.4	303	3	US-08-876-398A-38	Sequence 38, Appl
549	4	44.4	298	2	US-08-462-738-4	Sequence 4, Appli	622	4	44.4	303	3	US-08-876-398A-40	Sequence 40, Appl
550	4	44.4	299	3	US-08-872-437-2	Sequence 2, Appli	623	4	44.4	303	3	US-08-876-398A-42	Sequence 42, Appl
551	4	44.4	299	2	US-08-651-136C-12	Sequence 12, Appl	624	4	44.4	303	3	US-08-876-398A-44	Sequence 44, Appl
552	4	44.4	300	1	US-07-640-029-5	Sequence 5, Appli	625	4	44.4	303	3	US-08-876-398A-46	Sequence 46, Appl
553	4	44.4	300	3	US-08-705-875A-4	Sequence 4, Appli	626	4	44.4	303	3	US-08-876-398A-48	Sequence 48, Appl
554	4	44.4	300	3	US-08-705-875A-6	Sequence 6, Appli	627	4	44.4	303	3	US-08-876-398A-50	Sequence 50, Appl
555	4	44.4	300	4	US-09-220-731-21	Sequence 21, Appl	628	4	44.4	303	3	US-08-876-398A-52	Sequence 52, Appl
556	4	44.4	300	4	US-08-439-992A-5	Sequence 5, Appli	629	4	44.4	303	3	US-08-876-398A-54	Sequence 54, Appl
557	4	44.4	301	2	US-08-355-844-2	Sequence 2, Appli	630	4	44.4	303	3	US-08-876-398A-56	Sequence 56, Appl
558	4	44.4	301	5	PCT-US95-16126-2	Sequence 2, Appli	631	4	44.4	303	3	US-08-876-398A-58	Sequence 58, Appl
559	4	44.4	302	1	US-07-640-029-6	Sequence 6, Appli	632	4	44.4	303	3	US-08-876-398A-60	Sequence 60, Appl
560	4	44.4	302	1	US-07-921-807B-7	Sequence 7, Appli	633	4	44.4	303	3	US-08-876-398A-62	Sequence 62, Appl
561	4	44.4	302	1	US-07-921-807B-8	Sequence 8, Appli	634	4	44.4	303	3	US-08-876-398A-64	Sequence 64, Appl
562	4	44.4	302	1	US-08-441-944A-7	Sequence 7, Appli	635	4	44.4	303	3	US-08-876-398A-66	Sequence 66, Appl
563	4	44.4	302	1	US-08-441-944A-8	Sequence 8, Appli	636	4	44.4	303	3	US-08-876-398A-68	Sequence 68, Appl
564	4	44.4	302	4	US-08-439-992A-6	Sequence 6, Appli	637	4	44.4	304	1	US-08-900-711-5	Sequence 5, Appli
565	4	44.4	303	1	US-07-917-111-2	Sequence 3, Appli	638	4	44.4	304	2	US-08-815-336-2	Sequence 2, Appli
566	4	44.4	303	1	US-07-917-111-3	Sequence 3, Appli	639	4	44.4	304	2	US-08-415-343B-5	Sequence 5, Appli
567	4	44.4	303	1	US-08-479-638-2	Sequence 2, Appli	640	4	44.4	304	2	US-08-762-433-1	Sequence 1, Appli
568	4	44.4	303	1	US-08-479-638-3	Sequence 3, Appli	641	4	44.4	304	2	US-09-001-219-1	Sequence 1, Appli
569	4	44.4	303	2	US-08-294-871A-2	Sequence 2, Appli	642	4	44.4	305	1	US-08-348-792-6	Sequence 6, Appli
570	4	44.4	303	2	US-08-294-871A-4	Sequence 4, Appli	643	4	44.4	305	2	US-08-853-659A-47	Sequence 47, Appl
571	4	44.4	303	2	US-08-294-871A-6	Sequence 6, Appli	644	4	44.4	305	2	US-08-462-728-6	Sequence 6, Appli
572	4	44.4	303	2	US-08-294-871A-8	Sequence 8, Appli	645	4	44.4	306	1	US-08-454-186-11	Sequence 11, Appl
573	4	44.4	303	2	US-08-294-871A-10	Sequence 10, Appl	646	4	44.4	306	2	US-08-286-819A-33	Sequence 33, Appl
574	4	44.4	303	2	US-08-294-871A-12	Sequence 12, Appl	647	4	44.4	306	3	US-08-980-373-33	Sequence 33, Appl
575	4	44.4	303	2	US-08-294-871A-14	Sequence 14, Appl	648	4	44.4	306	3	US-09-064-033-11	Sequence 11, Appl
576	4	44.4	303	2	US-08-294-871A-16	Sequence 16, Appl	649	4	44.4	307	1	US-08-348-792-8	Sequence 8, Appli
577	4	44.4	303	2	US-08-294-871A-18	Sequence 18, Appl	650	4	44.4	307	2	US-08-462-738-8	Sequence 8, Appli
578	4	44.4	303	2	US-08-294-871A-20	Sequence 20, Appl	651	4	44.4	307	4	US-09-177-249-13	Sequence 13, Appl
579	4	44.4	303	2	US-08-294-871A-22	Sequence 22, Appl	652	4	44.4	308	3	US-08-705-875A-10	Sequence 10, Appl
580	4	44.4	303	2	US-08-294-871A-24	Sequence 24, Appl	653	4	44.4	311	2	US-08-602-359A-41	Sequence 41, Appl
581	4	44.4	303	2	US-08-294-871A-26	Sequence 26, Appl	654	4	44.4	312	1	US-08-094-128A-27	Sequence 27, Appl
582	4	44.4	303	2	US-08-294-871A-28	Sequence 28, Appl	655	4	44.4	312	1	US-08-455-674-27	Sequence 27, Appl
583	4	44.4	303	2	US-08-294-871A-30	Sequence 30, Appl	656	4	44.4	312	1	US-08-455-992-27	Sequence 27, Appl
584	4	44.4	303	2	US-08-294-871A-32	Sequence 32, Appl	657	4	44.4	312	1	US-08-455-992-27	Sequence 27, Appl
585	4	44.4	303	2	US-08-294-871A-34	Sequence 34, Appl	658	4	44.4	312	1	US-08-455-992-27	Sequence 27, Appl
586	4	44.4	303	2	US-08-294-871A-36	Sequence 36, Appl	659	4	44.4	312	2	US-09-031-485-2	Sequence 2, Appli
587	4	44.4	303	2	US-08-294-871A-38	Sequence 38, Appl	660	4	44.4	312	2	US-08-847-429A-2	Sequence 2, Appli
588	4	44.4	303	2	US-08-294-871A-40	Sequence 40, Appl	661	4	44.4	312	3	US-08-894-017-10	Sequence 10, Appl
589	4	44.4	303	2	US-08-294-871A-42	Sequence 42, Appl	662	4	44.4	312	3	US-09-065-474-2	Sequence 2, Appli
590	4	44.4	303	2	US-08-294-871A-44	Sequence 44, Appl	663	4	44.4	312	3	US-09-084-813-6	Sequence 6, Appli
591	4	44.4	303	2	US-08-294-871A-46	Sequence 46, Appl	664	4	44.4	312	4	US-09-216-295-21	Sequence 21, Appl
592	4	44.4	303	2	US-08-294-871A-48	Sequence 48, Appl	665	4	44.4	312	5	PCT-US92-00652-27	Sequence 27, Appl
593	4	44.4	303	2	US-08-294-871A-50	Sequence 50, Appl	666	4	44.4	312	5	PCT-US92-09662-6	Sequence 6, Appli
594	4	44.4	303	2	US-08-294-871A-52	Sequence 52, Appl	667	4	44.4	315	4	US-09-434-774-16	Sequence 16, Appl
595	4	44.4	303	2	US-08-294-871A-54	Sequence 54, Appl	668	4	44.4	317	1	US-07-841-646-11	Sequence 11, Appl
596	4	44.4	303	2	US-08-294-871A-56	Sequence 56, Appl	669	4	44.4	317	1	US-07-841-646-11	Sequence 11, Appl
597	4	44.4	303	2	US-08-294-871A-58	Sequence 58, Appl	670	4	44.4	317	1	US-07-709-949-2	Sequence 2, Appli
598	4	44.4	303	2	US-08-294-871A-60	Sequence 60, Appl	671	4	44.4	317	1	US-08-447-570-11	Sequence 11, Appl
599	4	44.4	303	2	US-08-294-871A-62	Sequence 62, Appl	672	4	44.4	317	2	US-08-449-700-11	Sequence 11, Appl
600	4	44.4	303	2	US-08-294-871A-64	Sequence 64, Appl	673	4	44.4	317	2	US-08-449-699A-11	Sequence 11, Appl
601	4	44.4	303	2	US-08-294-871A-66	Sequence 66, Appl	674	4	44.4	318	4	US-08-878-474-3	Sequence 4, Appli
602	4	44.4	303	2	US-08-294-871A-68	Sequence 68, Appl	675	4	44.4	319	4	US-08-836-075A-44	Sequence 44, Appl
603	4	44.4	303	3	US-08-876-398A-2	Sequence 2, Appli	676	4	44.4	323	2	US-08-435-149-2	Sequence 2, Appli
604	4	44.4	303	3	US-08-876-398A-4	Sequence 4, Appli	677	4	44.4	323	4	US-09-172-063-21	Sequence 21, Appl
605	4	44.4	303	3	US-08-876-398A-6	Sequence 6, Appli	678	4	44.4	323	4	US-09-172-063-22	Sequence 22, Appl
606	4	44.4	303	3	US-08-876-398A-8	Sequence 8, Appli	679	4	44.4	323	4	US-08-643-212-18	Sequence 18, Appl
607	4	44.4	303	3	US-08-876-398A-10	Sequence 10, Appl	680	4	44.4	323	4	US-08-643-212-20	Sequence 20, Appl
608	4	44.4	303	3	US-08-876-398A-12	Sequence 12, Appl	681	4	44.4	323	4	US-08-643-212-22	Sequence 22, Appl
609	4	44.4	303	3	US-08-876-398A-14	Sequence 14, Appl	682	4	44.4	323	4	US-08-643-212-24	Sequence 24, Appl
610	4	44.4	303	3	US-08-876-398A-16	Sequence 16, Appl	683	4	44.4	323	4	US-08-643-212-26	Sequence 26, Appl
611	4	44.4	303	3	US-08-876-398A-18	Sequence 18, Appl	684	4	44.4	323	4	US-08-643-212-28	Sequence 28, Appl

685	4	44.4	324	1	US-08-310-416A-14	Sequence 14, Appl	758	4	44.4	367	2	US-08-737-045-14	Sequence 14, Appl
686	4	44.4	324	2	US-08-888-171-14	Sequence 14, Appl	759	4	44.4	367	3	US-08-932-871B-2	Sequence 2, Appl
687	4	44.4	325	1	US-08-447-500-8	Sequence 8, Appl	760	4	44.4	367	3	US-09-476-819-2	Sequence 2, Appl
688	4	44.4	325	1	US-08-454-097-8	Sequence 8, Appl	761	4	44.4	368	4	US-08-961-083-76	Sequence 76, Appl
689	4	44.4	325	1	US-08-453-866-8	Sequence 8, Appl	762	4	44.4	368	4	US-09-413-574-4	Sequence 4, Appl
690	4	44.4	325	3	US-08-185-359-8	Sequence 8, Appl	763	4	44.4	368	4	US-08-991-677-6	Sequence 6, Appl
691	4	44.4	326	1	US-07-603-133B-28	Sequence 28, Appl	764	4	44.4	369	4	US-09-108-020-40	Sequence 40, Appl
692	4	44.4	326	2	US-08-786-606-8	Sequence 8, Appl	765	4	44.4	370	2	US-08-878-989-19	Sequence 19, Appl
693	4	44.4	328	1	US-08-414-926A-9	Sequence 9, Appl	766	4	44.4	370	3	US-08-369-826C-2	Sequence 2, Appl
694	4	44.4	328	2	US-08-926-922-9	Sequence 9, Appl	767	4	44.4	370	3	US-08-582-776C-2	Sequence 2, Appl
695	4	44.4	328	2	US-08-956-012-1	Sequence 1, Appl	768	4	44.4	370	3	US-09-248-835-26	Sequence 26, Appl
696	4	44.4	328	3	US-09-253-682-9	Sequence 9, Appl	769	4	44.4	370	3	US-08-434-831B-2	Sequence 2, Appl
697	4	44.4	328	6	5212074-4	Patent No. 5212074	770	4	44.4	370	4	US-09-272-796-19	Sequence 19, Appl
698	4	44.4	331	2	US-08-986-217-7	Sequence 7, Appl	771	4	44.4	371	4	US-09-104-308-1	Sequence 1, Appl
699	4	44.4	332	2	US-08-637-763B-6	Sequence 6, Appl	772	4	44.4	372	1	US-08-196-618-33	Sequence 33, Appl
700	4	44.4	332	2	US-08-637-763B-8	Sequence 8, Appl	773	4	44.4	372	1	US-08-681-953-33	Sequence 33, Appl
701	4	44.4	332	3	US-09-170-354-6	Sequence 6, Appl	774	4	44.4	373	2	US-08-912-129A-52	Sequence 52, Appl
702	4	44.4	332	3	US-09-170-354-8	Sequence 8, Appl	775	4	44.4	374	1	US-07-914-281-11	Sequence 11, Appl
703	4	44.4	333	2	US-08-869-506-3	Sequence 3, Appl	776	4	44.4	374	1	US-08-393-246-11	Sequence 11, Appl
704	4	44.4	333	3	US-09-128-967-3	Sequence 3, Appl	777	4	44.4	374	1	US-08-525-058A-11	Sequence 11, Appl
705	4	44.4	335	1	US-08-348-792-2	Sequence 2, Appl	778	4	44.4	374	2	US-08-696-731-11	Sequence 11, Appl
706	4	44.4	335	2	US-08-462-738-2	Sequence 2, Appl	779	4	44.4	374	4	US-09-042-531-11	Sequence 11, Appl
707	4	44.4	336	2	US-08-474-379C-88	Sequence 88, Appl	780	4	44.4	375	1	US-07-803-622E-7	Sequence 7, Appl
708	4	44.4	336	2	US-08-235-836C-52	Sequence 52, Appl	781	4	44.4	375	1	US-07-803-622E-9	Sequence 9, Appl
709	4	44.4	337	1	US-08-447-500-10	Sequence 10, Appl	782	4	44.4	378	2	US-09-055-097-1	Sequence 1, Appl
710	4	44.4	337	1	US-08-454-097-12	Sequence 12, Appl	783	4	44.4	379	1	US-08-347-029-2	Sequence 2, Appl
711	4	44.4	337	1	US-08-453-866-10	Sequence 10, Appl	784	4	44.4	379	1	US-08-490-803-2	Sequence 2, Appl
712	4	44.4	337	3	US-09-100-664A-11	Sequence 11, Appl	785	4	44.4	379	1	US-08-484-246-2	Sequence 2, Appl
713	4	44.4	337	3	US-09-100-664A-12	Sequence 12, Appl	786	4	44.4	379	1	US-08-457-254-2	Sequence 2, Appl
714	4	44.4	337	3	US-08-185-359-12	Sequence 12, Appl	787	4	44.4	379	2	US-08-959-927-4	Sequence 4, Appl
715	4	44.4	339	1	US-08-248-629A-5	Sequence 5, Appl	788	4	44.4	379	3	PCT-US94-08806-2	Sequence 2, Appl
716	4	44.4	339	1	US-08-451-932-5	Sequence 5, Appl	789	4	44.4	379	5	PCT-US94-08806-2	Sequence 2, Appl
717	4	44.4	339	1	US-08-452-260-5	Sequence 5, Appl	790	4	44.4	379	5	PCT-US95-01775-2	Sequence 2, Appl
718	4	44.4	339	1	US-08-326-785-5	Sequence 5, Appl	791	4	44.4	379	5	PCT-US95-01829-4	Sequence 4, Appl
719	4	44.4	339	2	US-08-612-788-5	Sequence 5, Appl	792	4	44.4	379	5	PCT-US95-14937-2	Sequence 2, Appl
720	4	44.4	339	2	US-08-605-598B-5	Sequence 5, Appl	793	4	44.4	379	5	PCT-US95-16626-2	Sequence 2, Appl
721	4	44.4	339	2	US-08-429-743-5	Sequence 5, Appl	794	4	44.4	382	4	US-09-142-551A-2	Sequence 2, Appl
722	4	44.4	339	2	US-08-866-735-5	Sequence 5, Appl	795	4	44.4	382	4	US-09-277-716-22	Sequence 22, Appl
723	4	44.4	339	5	US-09-066-028-5	Sequence 5, Appl	796	4	44.4	384	1	US-07-688-352C-40	Sequence 40, Appl
724	4	44.4	339	5	PCT-US95-05107-5	Sequence 5, Appl	797	4	44.4	384	2	US-08-474-379C-40	Sequence 40, Appl
725	4	44.4	341	4	US-09-196-293-14	Sequence 14, Appl	798	4	44.4	384	3	US-09-146-249E-40	Sequence 40, Appl
726	4	44.4	341	4	US-08-209-603E-14	Sequence 14, Appl	799	4	44.4	384	3	US-08-206-188B-40	Sequence 40, Appl
727	4	44.4	344	3	US-08-858-003-33	Sequence 33, Appl	800	4	44.4	386	6	5171840-5	Patent No. 5171840
728	4	44.4	344	3	US-09-078-166-33	Sequence 33, Appl	801	4	44.4	386	6	5480796-5	Patent No. 5480796
729	4	44.4	344	4	US-08-997-467-33	Sequence 33, Appl	802	4	44.4	387	1	US-08-314-596-40	Sequence 40, Appl
730	4	44.4	344	6	5171840-7	Patent No. 5171840	803	4	44.4	387	1	US-08-314-596-42	Sequence 42, Appl
731	4	44.4	344	6	5480796-7	Patent No. 5480796	804	4	44.4	387	1	US-08-320-982-40	Sequence 40, Appl
732	4	44.4	345	1	US-08-102-942A-2	Sequence 2, Appl	805	4	44.4	387	1	US-08-320-982-42	Sequence 42, Appl
733	4	44.4	346	2	US-08-602-359A-34	Sequence 34, Appl	806	4	44.4	387	2	US-08-872-302-6	Sequence 6, Appl
734	4	44.4	349	2	US-08-483-926A-12	Sequence 12, Appl	807	4	44.4	387	3	US-08-819-037-40	Sequence 40, Appl
735	4	44.4	351	5	PCT-US93-05703-2	Sequence 2, Appl	808	4	44.4	387	3	US-08-819-037-42	Sequence 42, Appl
736	4	44.4	352	2	US-08-483-926A-11	Sequence 11, Appl	809	4	44.4	390	1	US-08-347-792-15	Sequence 15, Appl
737	4	44.4	352	2	US-08-737-045-12	Sequence 12, Appl	810	4	44.4	390	1	US-08-431-357-15	Sequence 15, Appl
738	4	44.4	352	3	US-08-921-887-52	Sequence 52, Appl	811	4	44.4	390	4	US-08-392-542-3	Sequence 3, Appl
739	4	44.4	356	2	US-08-602-713-12	Sequence 12, Appl	812	4	44.4	390	4	US-08-894-327-3	Sequence 3, Appl
740	4	44.4	356	2	US-08-903-851-2	Sequence 2, Appl	813	4	44.4	390	5	PCT-US95-15353-15	Sequence 15, Appl
741	4	44.4	356	4	US-08-989-493-12	Sequence 12, Appl	814	4	44.4	393	1	US-08-047-041A-25	Sequence 25, Appl
742	4	44.4	353	2	US-08-697-221-17	Sequence 17, Appl	815	4	44.4	393	1	US-08-047-041A-27	Sequence 27, Appl
743	4	44.4	353	2	US-08-697-221-18	Sequence 18, Appl	816	4	44.4	393	1	US-08-347-792-2	Sequence 2, Appl
744	4	44.4	353	2	US-08-697-221-19	Sequence 19, Appl	817	4	44.4	393	1	US-08-390-516C-6	Sequence 6, Appl
745	4	44.4	353	2	US-08-697-221-20	Sequence 20, Appl	818	4	44.4	393	1	US-08-390-516C-8	Sequence 8, Appl
746	4	44.4	353	2	US-08-697-221-21	Sequence 21, Appl	819	4	44.4	393	1	US-08-431-357-2	Sequence 2, Appl
747	4	44.4	353	2	US-08-697-221-22	Sequence 22, Appl	820	4	44.4	393	1	US-08-390-515A-6	Sequence 6, Appl
748	4	44.4	353	2	US-08-697-221-23	Sequence 23, Appl	821	4	44.4	393	1	US-08-390-515A-8	Sequence 8, Appl
749	4	44.4	353	2	US-08-697-221-24	Sequence 24, Appl	822	4	44.4	393	2	US-08-795-006A-32	Sequence 32, Appl
750	4	44.4	354	1	US-08-447-500-12	Sequence 12, Appl	823	4	44.4	393	2	US-08-697-221-2	Sequence 2, Appl
751	4	44.4	354	1	US-08-454-097-10	Sequence 10, Appl	824	4	44.4	393	2	US-08-697-221-3	Sequence 3, Appl
752	4	44.4	354	1	US-08-453-866-12	Sequence 12, Appl	825	4	44.4	393	2	US-08-697-221-4	Sequence 4, Appl
753	4	44.4	354	2	US-08-204-288-2	Sequence 2, Appl	826	4	44.4	393	2	US-08-697-221-11	Sequence 11, Appl
754	4	44.4	354	2	US-08-185-359-10	Sequence 10, Appl	827	4	44.4	393	2	US-08-697-221-12	Sequence 12, Appl
755	4	44.4	354	4	US-09-338-671-2	Sequence 2, Appl	828	4	44.4	393	2	US-08-697-221-13	Sequence 13, Appl
756	4	44.4	365	4	US-08-715-325-2	Sequence 2, Appl	829	4	44.4	393	2	US-08-697-221-14	Sequence 14, Appl
757	4	44.4	367	2	US-08-514-451A-7	Sequence 7, Appl	830	4	44.4	393	2	US-08-697-221-15	Sequence 15, Appl

831	4	44.4	393	2	US-08-697-221-16	Sequence 16, Appl	904	4	44.4	430	1	US-08-643-763A-19	Sequence 19, Appl
832	4	44.4	393	2	US-08-697-221-25	Sequence 25, Appl	905	4	44.4	430	1	US-08-462-623-19	Sequence 19, Appl
833	4	44.4	393	2	US-08-697-221-26	Sequence 26, Appl	906	4	44.4	430	1	US-08-451-953A-19	Sequence 19, Appl
834	4	44.4	393	2	US-08-697-221-27	Sequence 27, Appl	907	4	44.4	430	2	US-08-459-346-4	Sequence 4, Appl
835	4	44.4	393	2	US-08-697-221-28	Sequence 28, Appl	908	4	44.4	430	2	US-08-445-468A-19	Sequence 19, Appl
836	4	44.4	393	2	US-08-801-718-6	Sequence 6, Appl	909	4	44.4	430	2	US-08-901-200A-6	Sequence 6, Appl
837	4	44.4	393	2	US-08-801-718-8	Sequence 8, Appl	910	4	44.4	430	2	US-08-449-700-25	Sequence 25, Appl
838	4	44.4	393	2	US-08-247-904B-12	Sequence 12, Appl	911	4	44.4	430	2	US-08-449-699A-25	Sequence 25, Appl
839	4	44.4	393	3	US-08-767-942A-23	Sequence 23, Appl	912	4	44.4	430	2	US-08-461-397A-19	Sequence 19, Appl
840	4	44.4	393	4	US-08-392-542-2	Sequence 2, Appl	913	4	44.4	430	2	US-08-912-086-19	Sequence 19, Appl
841	4	44.4	393	4	US-09-184-073-32	Sequence 32, Appl	914	4	44.4	430	3	US-08-278-730A-19	Sequence 19, Appl
842	4	44.4	393	4	US-08-328-673A-9	Sequence 9, Appl	915	4	44.4	430	3	US-08-889-419-4	Sequence 4, Appl
843	4	44.4	393	4	US-08-894-327-2	Sequence 2, Appl	916	4	44.4	430	3	US-08-445-467-19	Sequence 19, Appl
844	4	44.4	393	5	PCT-US95-15353-2	Sequence 2, Appl	917	4	44.4	430	3	US-08-480-515A-19	Sequence 19, Appl
845	4	44.4	394	3	US-08-466-368-2	Sequence 2, Appl	918	4	44.4	430	4	US-09-219-391-6	Sequence 6, Appl
846	4	44.4	394	6	5223418-2	Patent No. 5223418	919	4	44.4	430	4	PCT-US90-07654-2	Sequence 2, Appl
847	4	44.4	396	2	US-08-926-253-2	Sequence 2, Appl	920	4	44.4	430	5	PCT-US92-01968-19	Sequence 19, Appl
848	4	44.4	396	3	US-09-120-074-2	Sequence 2, Appl	921	4	44.4	430	5	PCT-US93-05446-9	Sequence 9, Appl
849	4	44.4	396	4	US-09-330-611-12	Sequence 12, Appl	922	4	44.4	430	5	PCT-US93-07189-4	Sequence 4, Appl
850	4	44.4	398	1	US-08-446-777-2	Sequence 2, Appl	923	4	44.4	430	5	PCT-US93-07190-19	Sequence 19, Appl
851	4	44.4	398	2	US-08-599-171A-29	Sequence 29, Appl	924	4	44.4	430	5	PCT-US93-07221-19	Sequence 19, Appl
852	4	44.4	398	2	US-08-284-391B-29	Sequence 29, Appl	925	4	44.4	430	5	PCT-US93-08742-19	Sequence 19, Appl
853	4	44.4	398	2	US-08-646-590B-29	Sequence 29, Appl	926	4	44.4	430	5	PCT-US93-08808-19	Sequence 19, Appl
854	4	44.4	398	3	US-09-069-226-29	Sequence 29, Appl	927	4	44.4	430	5	PCT-US93-08885-19	Sequence 19, Appl
855	4	44.4	398	4	US-09-412-184-29	Sequence 29, Appl	928	4	44.4	430	5	PCT-US93-10520-6	Sequence 6, Appl
856	4	44.4	399	2	US-08-839-581A-2	Sequence 2, Appl	929	4	44.4	431	1	US-07-841-646-2	Sequence 2, Appl
857	4	44.4	400	1	US-09-023-591A-2	Sequence 2, Appl	930	4	44.4	431	1	US-07-901-703-2	Sequence 2, Appl
858	4	44.4	400	1	US-07-764-731B-10	Sequence 10, Appl	931	4	44.4	431	1	US-07-539-736-4	Sequence 4, Appl
859	4	44.4	401	2	US-08-591-079-4	Sequence 4, Appl	932	4	44.4	431	1	US-08-147-023-2	Sequence 2, Appl
860	4	44.4	401	2	US-08-591-079-6	Sequence 6, Appl	933	4	44.4	431	1	US-08-206-864-2	Sequence 2, Appl
861	4	44.4	405	1	US-08-370-193A-9	Sequence 9, Appl	934	4	44.4	431	1	US-08-278-729A-17	Sequence 17, Appl
862	4	44.4	405	1	US-09-413-574-2	Sequence 2, Appl	935	4	44.4	431	1	US-08-480-528A-4	Sequence 4, Appl
863	4	44.4	406	1	US-07-978-687-38	Sequence 38, Appl	936	4	44.4	431	1	US-08-479-666-4	Sequence 4, Appl
864	4	44.4	406	5	PCT-US91-05801-38	Sequence 38, Appl	937	4	44.4	431	1	US-08-155-343A-17	Sequence 17, Appl
865	4	44.4	408	1	US-07-841-646-15	Sequence 15, Appl	938	4	44.4	431	1	US-08-406-672-17	Sequence 17, Appl
866	4	44.4	408	1	US-08-147-023-15	Sequence 15, Appl	939	4	44.4	431	1	US-08-643-563A-17	Sequence 17, Appl
867	4	44.4	408	1	US-08-447-570-15	Sequence 15, Appl	940	4	44.4	431	1	US-08-447-570-2	Sequence 2, Appl
868	4	44.4	408	2	US-08-449-700-15	Sequence 15, Appl	941	4	44.4	431	1	US-08-643-763A-17	Sequence 17, Appl
869	4	44.4	408	2	US-08-449-699A-15	Sequence 15, Appl	942	4	44.4	431	1	US-08-462-623-17	Sequence 17, Appl
870	4	44.4	409	2	US-08-809-740A-5	Sequence 5, Appl	943	4	44.4	431	1	US-08-451-953A-17	Sequence 17, Appl
871	4	44.4	410	1	US-08-123-345A-7	Sequence 7, Appl	944	4	44.4	431	2	US-08-459-346-2	Sequence 2, Appl
872	4	44.4	410	1	US-08-123-345A-7	Sequence 7, Appl	945	4	44.4	431	2	US-08-445-468A-17	Sequence 17, Appl
873	4	44.4	410	1	US-08-073-807A-16	Sequence 16, Appl	946	4	44.4	431	2	US-08-901-200A-4	Sequence 4, Appl
874	4	44.4	410	2	US-08-472-666-2	Sequence 2, Appl	947	4	44.4	431	2	US-08-481-337A-10	Sequence 10, Appl
875	4	44.4	410	4	US-09-431-573-4	Sequence 4, Appl	948	4	44.4	431	2	US-08-449-700-2	Sequence 2, Appl
876	4	44.4	410	4	US-09-431-573-5	Sequence 5, Appl	949	4	44.4	431	2	US-07-989-847-6	Sequence 6, Appl
877	4	44.4	410	5	PCT-US96-07615-2	Sequence 2, Appl	950	4	44.4	431	2	US-08-449-699A-2	Sequence 2, Appl
878	4	44.4	413	2	US-08-481-814A-8	Sequence 8, Appl	951	4	44.4	431	2	US-08-696-268B-4	Sequence 4, Appl
879	4	44.4	413	2	US-08-282-197C-49	Sequence 49, Appl	952	4	44.4	431	2	US-08-461-397A-17	Sequence 17, Appl
880	4	44.4	413	3	US-08-836-582-2	Sequence 2, Appl	953	4	44.4	431	2	US-08-912-086-17	Sequence 17, Appl
881	4	44.4	415	1	US-08-381-936-2	Sequence 2, Appl	954	4	44.4	431	3	US-08-278-730A-17	Sequence 17, Appl
882	4	44.4	415	3	US-08-943-374-2	Sequence 2, Appl	955	4	44.4	431	3	US-08-458-811-2	Sequence 2, Appl
883	4	44.4	415	4	US-08-927-219-8	Sequence 8, Appl	956	4	44.4	431	3	US-08-889-419-2	Sequence 2, Appl
884	4	44.4	417	2	US-08-672-564-9	Sequence 9, Appl	957	4	44.4	431	3	US-08-445-467-17	Sequence 17, Appl
885	4	44.4	418	5	PCT-US94-01321-72	Sequence 72, Appl	958	4	44.4	431	3	US-08-480-515A-17	Sequence 17, Appl
886	4	44.4	420	5	US-07-700-526-1	Sequence 1, Appl	959	4	44.4	431	3	US-08-459-129-2	Sequence 2, Appl
887	4	44.4	420	5	PCT-US92-03132-1	Sequence 1, Appl	960	4	44.4	431	4	US-09-219-391-4	Sequence 4, Appl
888	4	44.4	425	4	US-08-867-611-22	Sequence 22, Appl	961	4	44.4	431	4	US-08-469-411-6	Sequence 6, Appl
889	4	44.4	425	5	PCT-US92-06965A-27	Sequence 27, Appl	962	4	44.4	431	5	PCT-US90-07654-4	Sequence 4, Appl
890	4	44.4	426	1	US-08-455-550-21	Sequence 21, Appl	963	4	44.4	431	5	PCT-US91-07635-2	Sequence 2, Appl
891	4	44.4	429	1	US-08-234-783-4	Sequence 4, Appl	964	4	44.4	431	5	PCT-US92-01968-17	Sequence 17, Appl
892	4	44.4	429	1	US-08-456-907-4	Sequence 4, Appl	965	4	44.4	431	5	PCT-US93-05446-2	Sequence 2, Appl
893	4	44.4	429	5	PCT-US95-05523-4	Sequence 4, Appl	966	4	44.4	431	5	PCT-US93-07189-2	Sequence 2, Appl
894	4	44.4	430	1	US-07-841-646-25	Sequence 25, Appl	967	4	44.4	431	5	PCT-US93-07190-17	Sequence 17, Appl
895	4	44.4	430	1	US-07-901-703-9	Sequence 9, Appl	968	4	44.4	431	5	PCT-US93-07231-17	Sequence 17, Appl
896	4	44.4	430	1	US-08-147-023-25	Sequence 25, Appl	969	4	44.4	431	5	PCT-US93-08742-17	Sequence 17, Appl
897	4	44.4	430	1	US-08-278-729A-19	Sequence 19, Appl	970	4	44.4	431	5	PCT-US93-08808-17	Sequence 17, Appl
898	4	44.4	430	1	US-08-480-528A-6	Sequence 6, Appl	971	4	44.4	431	5	PCT-US93-08885-17	Sequence 17, Appl
899	4	44.4	430	1	US-08-479-666-6	Sequence 6, Appl	972	4	44.4	431	5	PCT-US93-10520-4	Sequence 4, Appl
900	4	44.4	430	1	US-08-155-343A-19	Sequence 19, Appl	973	4	44.4	431	5	PCT-US95-05467-10	Sequence 10, Appl
901	4	44.4	430	1	US-08-406-672-19	Sequence 19, Appl	974	4	44.4	431	5	PCT-US95-06724-2	Sequence 2, Appl
902	4	44.4	430	1	US-08-643-563A-19	Sequence 19, Appl	975	4	44.4	432	5	US-08-411-607A-3	Sequence 3, Appl
903	4	44.4	430	1	US-08-447-570-25	Sequence 25, Appl	976	4	44.4	433	2	US-08-867-149-1	Sequence 1, Appl

```

577 4 44.4 433 2 US-08-808-374-1 Sequence 1, Appl1
978 4 44.4 433 3 US-09-100-409A-1 Sequence 1, Appl1
979 4 44.4 433 6 5171838-13 Patent No. 5171838
980 4 44.4 437 2 US-08-883-515-4 Sequence 4, Appl1
981 4 44.4 437 3 US-08-486-099-91 Sequence 91, Appl1
982 4 44.4 437 3 US-08-360-107A-101 Sequence 101, Appl1
983 4 44.4 437 3 US-08-484-223B-91 Sequence 91, Appl1
984 4 44.4 437 3 US-08-919-597-91 Sequence 91, Appl1
985 4 44.4 437 3 US-08-475-668A-91 Sequence 91, Appl1
986 4 44.4 437 3 US-08-485-551A-91 Sequence 91, Appl1
987 4 44.4 437 3 US-08-471-913A-91 Sequence 91, Appl1
988 4 44.4 437 4 US-08-485-264A-91 Sequence 91, Appl1
989 4 44.4 438 2 US-08-282-197C-59 Sequence 59, Appl1
990 4 44.4 439 1 US-08-553-999B-2 Sequence 2, Appl1
991 4 44.4 439 1 US-08-959-638-9 Sequence 9, Appl1
992 4 44.4 445 1 US-08-447-925-6 Sequence 6, Appl1
993 4 44.4 446 2 US-08-244-205-5 Sequence 5, Appl1
994 4 44.4 446 5 PCT-US92-10284-5 Sequence 5, Appl1
995 4 44.4 446 5 PCT-US94-01321-10 Sequence 10, Appl1
996 4 44.4 448 2 US-08-959-749-2 Sequence 2, Appl1
997 4 44.4 449 1 US-08-102-942A-4 Sequence 6, Appl1
998 4 44.4 449 1 US-08-102-942A-6 Sequence 18, Appl1
999 4 44.4 453 1 US-08-417-330A-18 Sequence 7, Appl1
1000 4 44.4 453 1 US-08-570-157-7 Sequence 7, Appl1

```

## ALIGNMENTS

```

RESULT 1
US-08-456-670B-26
; Sequence 26, Application US/08456670B
; Patent No. 5932415
; GENERAL INFORMATION:
; APPLICANT: SCHUBERT, PETER
; APPLICANT: NEUMANN, STEFRIED
; APPLICANT: PAMELZIK, MARTINA
; APPLICANT: LINKEWELER, WINFRIED
; APPLICANT: BURGER, CHRISTA
; APPLICANT: HOFMANN, GOTTFRIED
; APPLICANT: BUBERT, ANDREAS
; APPLICANT: GOEBEL, WERNER
; APPLICANT: KOHLER, STEFAN
; TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING
; TITLE OF INVENTION: LISTERIAS
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLER, WHITE, ZELANO & BRANTGAN, P.C.
; STREET: 2200 CLARENDON BLVD., SUITE 1400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456.670B
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: A35
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US 08/412.227
; FILING DATE: 27-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075.248
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4239567.4
; FILING DATE: 25-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4219111.4

```

```

; FILING DATE: 11-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: HAMLET-KING, DIANA
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: MERCK 1694D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; TELEX: 64191
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Listeria monocytogenes
; STRAIN: EGD
; US-08-456-670B-26

```

```

Query Match 100.0%; Score 9; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 VSTPVAPTO 9
Db 1 VSTPVAPTO 9

```

```

RESULT 2
US-08-127-499A-27
; Sequence 27, Application US/08127499A
; Patent No. 5510264
; GENERAL INFORMATION:
; APPLICANT: VAN ALSTYNE, Diane
; APPLICANT: SHARMA, Lawrence Rajendra
; TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED
; TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/127.499A
; FILING DATE: 28-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 51916/102/INBI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; US-08-127-499A-27

```

```

Query Match 100.0%; Score 9; DB 1; Length 21;

```

Best Local Similarity 100.0%; Pred. No. 0.001;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPTQ 9  
|||||  
DB 2 VSTPVAPTQ 10

## RESULT 3

US-08-482-847-27  
; Sequence 27, Application US/08482847  
; Patent No. 5556757

GENERAL INFORMATION:  
; APPLICANT: VAN ALSTYNE, Diane

APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPTIOTIC SITES FOR

TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 40

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington

STATE: D.C.  
COUNTRY: USA

ZIP: 20007-5109  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847

FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499

FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399

TELEX: 904136  
INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids

TYPE: amino acid  
STRANDEDNESS:

TOPOLOGY: unknown  
US-08-482-847-27

Query Match 100.0%; Score 9; DB 1; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.001;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPTQ 9  
|||||  
DB 2 VSTPVAPTQ 10

## RESULT 4

US-08-456-670B-39  
; Sequence 39, Application US/08456670B  
; Patent No. 5932415

GENERAL INFORMATION:  
; APPLICANT: SCHUBERT, PETER

APPLICANT: NEUMANN, STEGFRIED  
APPLICANT: PAMELZIK, MARTINA

APPLICANT: LINXWEILER, WINFRIED

APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED

APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER

APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING

TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLER, WHITE, ZELANO & BRANIGAN, P.C.

STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON

STATE: VIRGINIA  
COUNTRY: US

ZIP: 22201  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B

FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227

FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4

FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 421911.4  
FILING DATE: 11-JUN-1992

ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA

REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333

TELEFAX: 703-243-6410  
TELEX: 64191

INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:

LENGTH: 232 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: peptide

ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes

STRAIN: EGD  
US-08-456-670B-39

Query Match 100.0%; Score 9; DB 2; Length 232;  
Best Local Similarity 100.0%; Pred. No. 0.0089;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPVAPTQ 9  
|||||  
DB 76 VSTPVAPTQ 84

## RESULT 5

US-08-456-670B-40  
; Sequence 40, Application US/08456670B  
; Patent No. 5932415

GENERAL INFORMATION:  
; APPLICANT: SCHUBERT, PETER

APPLICANT: NEUMANN, STEGFRIED  
APPLICANT: PAMELZIK, MARTINA

APPLICANT: LINXWELLER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: BUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
TITLE OF INVENTION: LISTERIAS  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ. ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 478 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-40

Query Match 100.0%; Score 9; DB 2; Length 478;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTRVAPTQ 9  
Db 145 VSTRVAPTQ 153

RESULT 6  
US-08-127-499A-26  
Sequence 26, Application US/08127499A  
Patent No. 5510264  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra

TITLE OF INVENTION: ANTIBODIES WHICH BIND MENINGITIS RELATED  
TITLE OF INVENTION: HOMOLOGOUS ANTIGENIC SEQUENCES  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/127,499A  
FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/102/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ. ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-127-499A-26

Query Match 100.0%; Score 9; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTRVAPTQ 9  
Db 145 VSTRVAPTQ 153

RESULT 7  
US-08-482-847-26  
Sequence 26, Application US/08482847  
Patent No. 5556757  
GENERAL INFORMATION:  
APPLICANT: VAN ALSTYNE, Diane  
APPLICANT: SHARMA, Lawrence Rajendra  
TITLE OF INVENTION: PEPTIDES REPRESENTING EPITOPIC SITES FOR  
TITLE OF INVENTION: BACTERIAL AND VIRAL MENINGITIS CAUSING AGENTS AND THEIR  
TITLE OF INVENTION: CNS CARRIER, ANTIBODIES THEREOF, AND USES THEREOF  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,847  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/127,499



FILING DATE: 28-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 51916/104/INBI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
US-08-482-847-26

Query Match 100.0%; Score 9; DB 1; Length 484;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 STPVAPQTQ 9  
|||||  
Db 145 STPVAPQTQ 153

RESULT 8  
US-08-456-670B-30  
Sequence 30, Application us/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINWELTER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 421911.4  
FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:

NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694DI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6533  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-30

Query Match 88.9%; Score 8; DB 2; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0066;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 STPVAPQTQ 9  
|||||  
Db 1 STPVAPQTQ 8

RESULT 9  
US-08-456-670B-31  
Sequence 31, Application us/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINWELTER, WINFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 421911.4  
FILING DATE: 11-JUN-1992

ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
US-08-456-670B-31

Query Match 66.7%; Score 6; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.68;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PVAPQ 9  
11111  
Db 1 PVAPQ 6

RESULT 10  
US-08-456-670B-17  
Sequence 17, Application US/08456670B  
Patent No. 5932415  
GENERAL INFORMATION:  
APPLICANT: SCHUBERT, PETER  
APPLICANT: NEUMANN, SIEGFRIED  
APPLICANT: PAMELZIK, MARTINA  
APPLICANT: LINXWEILER, WILFRIED  
APPLICANT: BURGER, CHRISTA  
APPLICANT: HOFMANN, GOTTFRIED  
APPLICANT: HUBERT, ANDREAS  
APPLICANT: GOEBEL, WERNER  
APPLICANT: KOHLER, STEFAN  
TITLE OF INVENTION: PROCESSES AND AGENTS FOR DETECTING  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MILLER, WHITE, ZELANO & BRANTGAN, P.C.  
STREET: 2200 CLARENDON BLVD., SUITE 1400  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: US  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,670B  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/412,227  
FILING DATE: 27-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/075,248  
FILING DATE: 11-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4239567.4  
FILING DATE: 25-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 4219111.4

FILING DATE: 11-JUN-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: HAMLET-KING, DIANA  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: MERCK 1694D1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
ORGANISM: Listeria monocytogenes  
STRAIN: EGD  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 1-7  
OTHER INFORMATION: /note= "SOME OR ALL Xaa AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 14-20  
OTHER INFORMATION: /note= "SOME OR ALL Xaa AMINO ACIDS  
OTHER INFORMATION: MAY BE ABSENT"  
US-08-456-670B-17

Query Match 66.7%; Score 6; DB 2; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PVAPQ 9  
11111  
Db 8 PVAPQ 13

RESULT 11  
US-09-001-984C-55  
Sequence 55, Application US/09001984C  
Patent No. 6245331  
GENERAL INFORMATION:  
APPLICANT: Laal, Suman  
APPLICANT: Zolla-Pazner, Susan  
TITLE OF INVENTION: EARLY DETECTION OF MYCOBACTERIAL DISEASE  
FILE REFERENCE: NYU-011  
CURRENT APPLICATION NUMBER: US/09/001,984C  
CURRENT FILING DATE: 1997-12-31  
PRIOR APPLICATION NUMBER: 60/034,003  
PRIOR FILING DATE: 1996-12-31  
NUMBER OF SEQ ID NOS: 106  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 55  
LENGTH: 11  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis strain H37rv  
US-09-001-984C-55

Query Match 55.6%; Score 5; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 8.1;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
11111  
Db 1 TPVAP 5

RESULT 12

US-07-899-535A-1  
; Sequence 1, Application US/07899535A  
; Patent No. 5428011  
; GENERAL INFORMATION:  
; APPLICANT: Sheth, Anil R.  
; APPLICANT: Garde, Seema  
; TITLE OF INVENTION: Pharmaceutical Preparations For  
; TITLE OF INVENTION: Inhibiting Tumours Associated With Prostate  
; TITLE OF INVENTION: Adenocarcinoma, Stomach Cancer and Breast Cancer.  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Mr. George Loud  
; STREET: 2001 Jefferson Davis Highway, Suite 306  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/899,535A  
; FILING DATE: 16-JUN-1992  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Loud, George A.  
; REGISTRATION NUMBER: 25,814  
; REFERENCE/DOCKET NUMBER: S&B-A835  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-415-0960  
; TELEFAX: 703-415-0962  
; TELEX: 24 8614  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 94 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEetical: NO  
; US-07-899-535A-1

Query Match 55.6%; Score 5; DB 1; Length 94;  
Best Local Similarity 100.0%; Pred. No. 57;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSTPV 5  
|11111  
DB 53 VSTPV 57

RESULT 13  
US-08-086-428B-57  
; Sequence 57, Application US/08086428B  
; Patent No. 5514539  
; GENERAL INFORMATION:  
; APPLICANT: BUKH, J., MILLER, R.H. AND  
; APPLICANT: PURCELL, R.H.  
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE  
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE  
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN  
; NUMBER OF SEQUENCES: 159  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN & FINNEGAN  
; STREET: 345 PARK AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK

COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/086,428B  
FILING DATE: 29-JUN-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4070  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
ORIGINAL SOURCE:  
ORGANISM: homosapiens  
INDIVIDUAL ISOLATE: S18  
US-08-086-428B-57

Query Match 55.6%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PVAPT 8  
|11111  
DB 50 PVAPT 54

RESULT 14  
US-08-440-103-44  
; Sequence 44, Application US/08440103  
; Patent No. 5670152  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, Amy J.  
; APPLICANT: Houghton, Michael  
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,103  
; FILING DATE: 12-MAY-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,368  
; FILING DATE:  
; APPLICATION NUMBER: US 07/759,575  
; FILING DATE: 13-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113

REFERENCE/DOCKET NUMBER: 0205.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2708  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-103-44

Query Match 55.6%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PVAPT 8  
Db 50 PVAPT 54

RESULT 15  
US-08-440-542-44  
Sequence 44, Application US/08440542  
Patent No. 5670153  
GENERAL INFORMATION:  
APPLICANT: Weiner, Amy J.  
TITLE OF INVENTION: Immunoactive Polypeptide Compositions  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,542  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,368  
FILING DATE:  
APPLICATION NUMBER: US 07/759,575  
FILING DATE: 13-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McClung, Barbara G.  
REGISTRATION NUMBER: 33,113  
REFERENCE/DOCKET NUMBER: 0205.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2708  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-542-44

Query Match 55.6%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PVAPT 8  
Db 50 PVAPT 54

RESULT 16  
US-08-231-368-44  
Sequence 44, Application US/08231368  
Patent No. 5756312  
GENERAL INFORMATION:  
APPLICANT: Weiner, Amy J.  
TITLE OF INVENTION: Immunoactive Polypeptide Compositions  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,368  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/759,575  
FILING DATE: 13-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McClung, Barbara G.  
REGISTRATION NUMBER: 33,113  
REFERENCE/DOCKET NUMBER: 0205.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2708  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-231-368-44

Query Match 55.6%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PVAPT 8  
Db 50 PVAPT 54

RESULT 17  
US-08-440-210-44  
Sequence 44, Application US/08440210  
Patent No. 5766845  
GENERAL INFORMATION:  
APPLICANT: Weiner, Amy J.  
TITLE OF INVENTION: Immunoactive Polypeptide Compositions  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville

STATE: CA  
COUNTRY: USA  
ZIP: 94608  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,210  
FILING DATE: 12-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/231,368  
FILING DATE: US 07/759,575  
APPLICATION NUMBER: US 07/759,575  
FILING DATE: 13-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McClung, Barbara G.  
REGISTRATION NUMBER: 33,113  
REFERENCE/DOCKET NUMBER: 0205.001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2708  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-440-210-44

Query Match 55.6%; Score 5; DB 1; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
DB 50 PVAPT 54

RESULT 18  
US-08-468-570-57  
Sequence 57, Application US/08468570  
Patent No. 5871962  
GENERAL INFORMATION:  
APPLICANT: BURK, J., MILLER, R.H. AND  
APPLICANT: PURCELL, R.H.  
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE  
TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE  
TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN  
TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES  
NUMBER OF SEQUENCES: 159  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,570  
FILING DATE: 6-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/086,428  
FILING DATE: 29-JUN-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4070US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
ORIGINAL SOURCE:  
ORGANISM: homosapiens  
INDIVIDUAL ISOLATE: S18  
US-08-468-570-57

Query Match 55.6%; Score 5; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
DB 50 PVAPT 54

RESULT 19  
US-08-290-665A-57  
Sequence 57, Application US/08290665A  
Patent No. 5882852  
GENERAL INFORMATION:  
APPLICANT: BURK, J., MILLER, R.H. AND  
APPLICANT: PURCELL, R.H.  
TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND  
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS  
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE  
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES  
NUMBER OF SEQUENCES: 263  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290,665A  
FILING DATE: 15-AUG-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid

STRANDEDNESS: unknown  
TOPOLOGY: unknown  
ORIGINAL SOURCE:  
ORGANISM: homosapiens  
INDIVIDUAL ISOLATE: S18  
US-08-290-665A-57

Query Match 55.6%; Score 5; DB 2; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
|||||  
DB 50 PVAPT 54

RESULT 20  
US-08-444-818-167  
Sequence 167, Application US/08444818  
Patent No. 6150087  
GENERAL INFORMATION:  
APPLICANT: Chien, David Y.  
APPLICANT: Rutter, William J.  
TITLE OF INVENTION: NANBV Diagnostics and Vaccines  
NUMBER OF SEQUENCES: 777  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
City: Emeryville  
STATE: CA  
COUNTRY: USA  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,818  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,590  
FILING DATE: 14-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Harbin, Ailsa A.  
REGISTRATION NUMBER: 33,895  
REFERENCE/DOCKET NUMBER: 0110.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (508)359-3876  
TELEFAX: (508)359-3885  
INFORMATION FOR SEQ ID NO: 167:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: JH27  
US-08-444-818-167

Query Match 55.6%; Score 5; DB 4; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
|||||  
DB 125 PVAPT 129

RESULT 21  
PCT-US95-10398-57  
Sequence 57, Application PC/TUS9510398  
GENERAL INFORMATION:  
APPLICANT: BUKH, J., MILLER, R. H. AND  
APPLICANT: PURCELL, R. H.

TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED  
TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND  
TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS  
TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE  
TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES  
NUMBER OF SEQUENCES: 263  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN  
STREET: 345 PARK AVENUE  
City: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/10398  
FILING DATE: 15-AUG-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/086,428  
FILING DATE: 29 JUNE 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/290/665  
FILING DATE: 15 AUGUST 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: RICHARD W. BORK  
REGISTRATION NUMBER: 36,459  
REFERENCE/DOCKET NUMBER: 2026-4116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
ORIGINAL SOURCE:  
ORGANISM: homosapiens  
INDIVIDUAL ISOLATE: S18  
PCT-US95-10398-57

Query Match 55.6%; Score 5; DB 5; Length 192;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
|||||  
DB 50 PVAPT 54

RESULT 22  
US-08-969-644-12  
Sequence 12, Application US/08969644  
Patent No. 6096519  
GENERAL INFORMATION:  
APPLICANT: Ratti, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA

TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM: RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/969,644  
FILING DATE: 13-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/467,152  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-969-644-12

Query Match 55.6%; Score 5; DB 3; Length 264;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
|||||  
DB 40 STPVA 44

RESULT 23  
US-08-444-189-12  
Sequence 12, Application US/08444189  
Patent No. 6110705  
GENERAL INFORMATION:  
APPLICANT: Rattl, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCID PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM: RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA

ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,189  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/180,528  
FILING DATE:  
APPLICATION NUMBER: US/07/991,512  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-444-189-12

Query Match 55.6%; Score 5; DB 3; Length 264;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
|||||  
DB 40 STPVA 44

RESULT 24  
US-08-465-465-2  
Sequence 2, Application US/08465465  
Patent No. 6210968  
GENERAL INFORMATION:  
APPLICANT: Giulio Rattl  
TITLE OF INVENTION: RECOMBINANT CHLAMYDIA TRACHOMATIS PCP3  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 6210968r1s  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch disk, 720 KB  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,465  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/229,980  
FILING DATE: 19-APR-1994

ATTORNEY/AGENT INFORMATION:  
NAME: Paul K. Legaard  
REGISTRATION NUMBER: 38,534  
REFERENCE/DOCKET NUMBER: CHIR-0039  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-568-3100  
TELEFAX: 215-568-3439  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-465-465-2

Query Match 55.6%; Score 5; DB 4; Length 264;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 STPVA 6  
|||||  
DB 40 STPVA 44

RESULT 25  
US-08-468-544-12  
Sequence 12, Application US/08468544  
Patent No. 6248563  
GENERAL INFORMATION:  
APPLICANT: Rattl, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM, RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,544  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/661,820  
FILING DATE: 28-FEB-1991  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 264 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-468-544-12

Query Match 55.6%; Score 5; DB 4; Length 264;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 STPVA 6  
|||||  
DB 40 STPVA 44

RESULT 26  
US-08-382-184-3  
Sequence 3, Application US/08382184  
Patent No. 5714593  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: MICROBACTERIAL PROTEINS,  
TITLE OF INVENTION: MICROORGANISMS PRODUCING THEM AND THEIR USE FOR VACCINES  
TITLE OF INVENTION: AND FOR THE DETECTION OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
ADDRESSEE: NEUSTADT, P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/382,184  
FILING DATE: 01-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 22640720  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 286 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
US-08-382-184-3

Query Match 55.6%; Score 5; DB 1; Length 286;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
|||||  
DB 27 TPVAP 31

RESULT 27  
US-08-641-356-3  
Sequence 3, Application US/08641356  
Patent No. 5865130  
GENERAL INFORMATION:  
APPLICANT:



;; TITLE OF INVENTION: MICROBACTERIAL PROTEINS,  
;; TITLE OF INVENTION: MICROORGANISMS PRODUCING THEM AND THEIR USE FOR VACCINES  
;; TITLE OF INVENTION: AND FOR THE DETECTION OF TUBERCULOSIS  
;; NUMBER OF SEQUENCES: 3  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
;; ADDRESSEE: NEUSTADT, P.C.  
;; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
;; CITY: Arlington  
;; STATE: Virginia  
;; COUNTRY: U.S.A.  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/641,356  
;; FILING DATE:  
;; CLASSIFICATION: 536  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/382,184  
;; FILING DATE: 01-FEB-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OBLON, NORMAN F.  
;; REGISTRATION NUMBER: 24,618  
;; REFERENCE/DOCKET NUMBER: 22640720  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-413-3000  
;; TELEFAX: 703-413-2220  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 286 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; HYPOTHEICAL: NO  
;; US-08-641-356-3

Query Match 55.6%; Score 5; DB 2; Length 286;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
Db 27 TPVAP 31

RESULT 28  
US-09-132-528-4  
; Sequence 4, Application US/09132528A  
; Patent No. 6221353  
; GENERAL INFORMATION:  
; APPLICANT: LAQUEYERIE, ANNE  
; APPLICANT: MARCHAL, GILLES  
; APPLICANT: PESCHER, PASCALE  
; APPLICANT: ROMAIN, FELIX  
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM  
; TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 0660-0138-0DIY  
; CURRENT APPLICATION NUMBER: US/09/132,528A  
; CURRENT FILING DATE: 1998-08-11  
; EARLIER APPLICATION NUMBER: 08/641,356  
; EARLIER FILING DATE: 1996-04-30  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 286  
; TYPE: PRT

;; ORGANISM: Mycobacterium tuberculosis  
;; US-09-132-528-4

Query Match 55.6%; Score 5; DB 4; Length 286;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
Db 27 TPVAP 31

RESULT 29  
US-08-875-494-3  
; Sequence 3, Application US/08875494  
; Patent No. 6221610  
; GENERAL INFORMATION:  
; APPLICANT: LAQUEYERIE, ANNE  
; APPLICANT: MARCHAL, GILLES  
; APPLICANT: PESCHER, PASCALE  
; APPLICANT: ROMAIN, FELIX  
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM  
; TITLE OF INVENTION: AND THEIR USES FOR VACCINES AND FOR THE DETECTION OF  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 0660-0122-0 PCT  
; CURRENT APPLICATION NUMBER: US/08/875,494  
; CURRENT FILING DATE: 1997-08-01  
; EARLIER APPLICATION NUMBER: PCT/FR96/00166  
; EARLIER FILING DATE: 1996-02-01  
; EARLIER APPLICATION NUMBER: 382184  
; EARLIER FILING DATE: 1995-02-01  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 286  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
;; US-08-875-494-3

Query Match 55.6%; Score 5; DB 4; Length 286;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
Db 27 TPVAP 31

RESULT 30  
US-08-382-184-2  
; Sequence 2, Application US/08382184  
; Patent No. 5714593  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: MICROBACTERIAL PROTEINS,  
; TITLE OF INVENTION: MICROORGANISMS PRODUCING THEM AND THEIR USE FOR VACCINES  
; TITLE OF INVENTION: AND FOR THE DETECTION OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
; ADDRESSEE: NEUSTADT, P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/382,184  
FILING DATE: 01-FEB-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 22640720  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 325 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-382-184-2

Query Match 55.6%; Score 5; DB 1; Length 325;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
|||||  
Db 66 TPVAP 70

RESULT 31  
US-08-641-356-2  
Sequence 2, Application US/08641356  
Patent No. 5866130  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: MICROBACTERIAL PROTEINS, THEM AND THEIR USE FOR VACCINES  
TITLE OF INVENTION: AND FOR THE DETECTION OF TUBERCULOSIS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &  
ADDRESS: NEUSTADT, P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/641,356  
FILING DATE:  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/382,184  
FILING DATE: 01-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 22640720  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 325 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-641-356-2

Query Match 55.6%; Score 5; DB 2; Length 325;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
|||||  
Db 66 TPVAP 70

RESULT 32  
US-09-132-528-2  
Sequence 2, Application US/09132528A  
Patent No. 6221353  
GENERAL INFORMATION:  
APPLICANT: LAQUEYERIE, Anne  
APPLICANT: MARCHAL, Gilles  
APPLICANT: PESCHER, Pascale  
APPLICANT: ROMAIN, Felix  
TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM  
TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF  
TITLE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 0660-0138-ODIV  
CURRENT APPLICATION NUMBER: US/09/132,528A  
CURRENT FILING DATE: 1998-08-11  
EARLIER APPLICATION NUMBER: 08/641,356  
EARLIER FILING DATE: 1996-04-30  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 2  
LENGTH: 325  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
US-09-132-528-2

Query Match 55.6%; Score 5; DB 4; Length 325;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
|||||  
Db 66 TPVAP 70

RESULT 33  
US-09-132-528-3  
Sequence 3, Application US/09132528A  
Patent No. 6221353  
GENERAL INFORMATION:  
APPLICANT: LAQUEYERIE, Anne  
APPLICANT: MARCHAL, Gilles  
APPLICANT: PESCHER, Pascale  
APPLICANT: ROMAIN, Felix  
TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM  
TITLE OF INVENTION: AND THEIR USE FOR VACCINES AND FOR THE DETECTION OF  
TITLE OF INVENTION: TUBERCULOSIS  
FILE REFERENCE: 0660-0138-ODIV  
CURRENT APPLICATION NUMBER: US/09/132,528A  
CURRENT FILING DATE: 1998-08-11  
EARLIER APPLICATION NUMBER: 08/641,356  
EARLIER FILING DATE: 1996-04-30  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 3  
LENGTH: 325  
TYPE: PRT  
ORGANISM: Mycobacterium tuberculosis  
US-09-132-528-3

Query Match 55.6%; Score 5; DB 4; Length 325;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
|||||

DB 66 TPVAP 70

RESULT 34

US-08-875-494-2  
; Sequence 2, Application US/08875494  
; Patent No. 6221610

; GENERAL INFORMATION:  
; APPLICANT: LAOUEYERIE, ANNE

; APPLICANT: MARCHEL, GILLES  
; APPLICANT: PESCHER, PASCALE

; APPLICANT: ROMAIN, FELIX  
; TITLE OF INVENTION: MYCOBACTERIAL PROTEINS, MICROORGANISMS PRODUCING THEM

; TITLE OF INVENTION: AND THEIR USES FOR VACCINES AND FOR THE DETECTION OF  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 0660-0122-0 PCT

; CURRENT APPLICATION NUMBER: US/08/875,494  
; CURRENT FILING DATE: 1997-08-01

; EARLIER APPLICATION NUMBER: PCT/FR96/00166  
; EARLIER FILING DATE: 1996-02-01

; EARLIER APPLICATION NUMBER: 382184  
; EARLIER FILING DATE: 1995-02-01

; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2  
; LENGTH: 325

; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis

US-08-875-494-2

Query Match 55.6%; Score 5; DB 4; Length 325;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TPVAP 7  
|||||

DB 66 TPVAP 70

RESULT 35

US-08-444-818-152  
; Sequence 152, Application US/08444818  
; Patent No. 6150087

; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.

; APPLICANT: Rutger, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines

; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street

; CITY: Emeryville  
; STATE: CA

; COUNTRY: USA  
; ZIP: 94608-2916

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:

; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Harbin, Alisa A.

REGISTRATION NUMBER: 33,895

REFERENCE/DOCKET NUMBER: 0110.002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (508)359-3876

TELEFAX: (508)359-3885

INFORMATION FOR SEQ ID NO: 152:

SEQUENCE CHARACTERISTICS:

LENGTH: 339 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-444-818-152

QY 4 PVAPT 8  
|||||

DB 113 PVAPT 117

RESULT 36

US-08-444-818-156  
; Sequence 156, Application US/08444818  
; Patent No. 6150087

; GENERAL INFORMATION:  
; APPLICANT: Chien, David Y.

; APPLICANT: Rutger, William J.  
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines

; NUMBER OF SEQUENCES: 777  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street

; CITY: Emeryville  
; STATE: CA

; COUNTRY: USA  
; ZIP: 94608-2916

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/444,818  
; FILING DATE:

; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/403,590  
; FILING DATE: 14-MAR-1995

; ATTORNEY/AGENT INFORMATION:  
; NAME: Harbin, Alisa A.

; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 0110.002

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (508)359-3876

; TELEFAX: (508)359-3885  
; INFORMATION FOR SEQ ID NO: 156:

SEQUENCE CHARACTERISTICS:

LENGTH: 339 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE: Human 27

INDIVIDUAL ISOLATE: US-08-444-818-156

Query Match 55.6%; Score 5; DB 4; Length 339;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 PVAPT 8  
11111  
Db 113 PVAPT 117

RESULT 37  
US-08-137-627-4  
; Sequence 4, Application US/08137627  
; Patent No. 5681737  
; GENERAL INFORMATION:  
; APPLICANT: Gee, Pauline  
; APPLICANT: Maion, Dorothy M.  
; APPLICANT: Ames, Bruce N.  
; TITLE OF INVENTION: A DETECTION SYSTEM FOR MUTAGENS THAT  
; TITLE OF INVENTION: IDENTIFIES MUTAGENIC CHANGES  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert,  
; ADDRESSEE: Attn: Karen S. Smith  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/137,627  
; FILING DATE: 15-OCT-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, Karen S.  
; REGISTRATION NUMBER: 31,426  
; REFERENCE/DOCKET NUMBER: A-58312/KSS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-137-627-4

Query Match 55.6%; Score 5; DB 1; Length 359;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
11111  
Db 249 STPVA 253

RESULT 38  
US-08-865-348-4  
; Sequence 4, Application US/08865348  
; Patent No. 5869258  
; GENERAL INFORMATION:  
; APPLICANT: Gee, Pauline  
; APPLICANT: Maion, Dorothy M.  
; APPLICANT: Ames, Bruce N.  
; TITLE OF INVENTION: A DETECTION SYSTEM FOR MUTAGENS THAT  
; TITLE OF INVENTION: ALSO IDENTIFIES MUTAGENIC CHANGES  
; NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert,  
; ADDRESSEE: Attn: Karen S. Smith  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/865,348  
; FILING DATE: 28-MAY-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/137,627  
; FILING DATE: 15-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, Karen S.  
; REGISTRATION NUMBER: 31,426  
; REFERENCE/DOCKET NUMBER: A-58312-1/KSS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 359 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-865-348-4

Query Match 55.6%; Score 5; DB 2; Length 359;  
Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
11111  
Db 249 STPVA 253

RESULT 39  
US-08-969-644-23  
; Sequence 23, Application US/08969644  
; Patent No. 6096519  
; GENERAL INFORMATION:  
; APPLICANT: Rattli, Giulio  
; APPLICANT: Comanducci, Maurizio  
; APPLICANT: Tecce, Mario F.  
; APPLICANT: Giuliani, Marzia M.  
; TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA  
; TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
; TITLE OF INVENTION: THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
; TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
; STREET: 301 N. Washington Street  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22046-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/969,644  
FILING DATE: 13-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/467,152  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 370 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
STRAIN: E. coli ATCC 68315  
IMMEDIATE SOURCE:  
CLONE: plasmid p03/GO/MC1  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..370  
OTHER INFORMATION: /label= polypeptide  
OTHER INFORMATION: /note= "polypeptide is a fusion protein of the  
OTHER INFORMATION: RNA polymerase from bacteriophage MS2 and the  
OTHER INFORMATION: protein encoded by the ORF3D gene of C.  
FEATURE:  
NAME/KEY: Region  
LOCATION: 107..370  
OTHER INFORMATION: /label= region  
OTHER INFORMATION: /note= "this portion of the fusion protein is the  
OTHER INFORMATION: protein encoded by the ORF3D gene."  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..106  
OTHER INFORMATION: /label= region  
OTHER INFORMATION: /note= "this portion of the fusion protein is a  
OTHER INFORMATION: fragment of the RNA polymerase gene from the  
OTHER INFORMATION: bacteriophage MS2."  
US-08-969-644-23  
Query Match 55.6%; Score 5; DB 3; Length 370;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 STPVA 6  
DB 146 STPVA 150  
RESULT 40  
US-08-444-189-23  
Sequence 23, Application US/08444189  
Patent No. 6110705  
GENERAL INFORMATION:  
APPLICANT: Ratelli, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tecce, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM; RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID

NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,189  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/180,528  
FILING DATE:  
APPLICATION NUMBER: US/07/991,512  
FILING DATE:  
APPLICATION NUMBER: US/07/661,820  
FILING DATE:  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 370 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
STRAIN: E. coli ATCC 68315  
IMMEDIATE SOURCE:  
CLONE: plasmid p03/GO/MC1  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..370  
OTHER INFORMATION: /label= polypeptide  
OTHER INFORMATION: /note= "polypeptide is a fusion protein of the  
OTHER INFORMATION: RNA-polymerase from bacteriophage MS2 and the  
OTHER INFORMATION: protein encoded by the ORF3D gene of C.  
FEATURE:  
NAME/KEY: Region  
LOCATION: 107..370  
OTHER INFORMATION: /label= region  
OTHER INFORMATION: /note= "this portion of the fusion protein is the  
OTHER INFORMATION: protein encoded by the ORF3D gene."  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..106  
OTHER INFORMATION: /label= region  
OTHER INFORMATION: /note= "this portion of the fusion protein is a  
OTHER INFORMATION: fragment of the RNA polymerase gene from the  
OTHER INFORMATION: bacteriophage MS2."  
US-08-444-189-23  
Query Match 55.6%; Score 5; DB 3; Length 370;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 STPVA 6  
DB 146 STPVA 150

Db 146 STPVA 150

RESULT 41  
US-08-468-544-23  
Sequence 23, Application US/08468544  
Patent No. 6248563  
GENERAL INFORMATION:  
APPLICANT: Ratcl, Giulio  
APPLICANT: Comanducci, Maurizio  
APPLICANT: Tece, Mario F.  
APPLICANT: Giuliani, Marzia M.  
TITLE OF INVENTION: PCTD PLASMID ISOLATED FROM CHLAMYDIA  
TITLE OF INVENTION: TRACHOMATIS SEROTYPE D, ITS GENES AND PROTEINS ENCODED BY  
TITLE OF INVENTION: THEM: RECOMBINANT PLASMIDS FOR THE EXPRESSION OF SAID  
TITLE OF INVENTION: GENES IN HETEROLOGOUS SYSTEMS, PREPARATION OF SAID  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH  
STREET: 301 N. Washington Street  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22046-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,544  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/661,820  
FILING DATE: 28-FEB-1991  
APPLICATION NUMBER: IT MI 91A000314  
FILING DATE: 07-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Svensson, Leonard R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 1267-202P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 370 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
ORIGINAL SOURCE:  
STRAIN: E. coli ATCC 68315  
IMMEDIATE SOURCE:  
CLONE: plasmid P03/GO/MC1  
FEATURE:  
NAME/KEY: Region  
LOCATION: 1..370  
OTHER INFORMATION: /label- polypeptide  
OTHER INFORMATION: /note- "polypeptide is a fusion protein of the  
OTHER INFORMATION: RNA-polymerase from bacteriophage MS2 and the  
OTHER INFORMATION: protein encoded by the ORF3D gene of C.  
FEATURE:  
NAME/KEY: Region  
LOCATION: 107..370  
OTHER INFORMATION: /label- region  
OTHER INFORMATION: /note- "this portion of the fusion protein is the  
OTHER INFORMATION: protein encoded by the ORF3D gene."  
NAME/KEY: Region  
LOCATION: 1..106  
OTHER INFORMATION: /label- region

OTHER INFORMATION: /note- "this portion of the fusion protein is a  
OTHER INFORMATION: fragment of the RNA polymerase gene from the  
OTHER INFORMATION: bacteriophage MS2."  
US-08-468-544-23

Query Match 55.6%; Score 5; DB 4; Length 370;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 2 STPVA 6  
Db 146 STPVA 150

RESULT 42  
US-08-853-659A-41  
Sequence 41, Application US/08853659A  
Patent No. 5925522  
GENERAL INFORMATION:  
APPLICANT: Wong, K.K.; Saifer, J.D.  
TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection  
TITLE OF INVENTION: Of A  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Paul W. Zimmerman  
ADDRESSEE: Intellectual Property Services  
ADDRESSEE: Battelle Memorial Institute  
ADDRESSEE: PNNL P.O. Box 999  
STREET: Washington Way  
CITY: Richland  
STATE: Washington  
COUNTRY: U.S.A.  
ZIP: 99352  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette 3.50", 1.44 MB storage  
COMPUTER: IBM PC/XT/AT  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Word Processor (WordPerfect 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/853,659A  
FILING DATE: Unknown  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: none  
FILING DATE: n/a  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 377 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-853-659A-41

Query Match 55.6%; Score 5; DB 2; Length 377;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 1 VSTPV 5  
Db 294 VSTPV 298

RESULT 43  
US-08-852-824-4  
Sequence 4, Application US/08852824C  
Patent No. 6060272  
GENERAL INFORMATION:  
APPLICANT: Li et al.  
TITLE OF INVENTION: Human G-protein Coupled Receptors  
FILE REFERENCE: 1488.1220000  
CURRENT APPLICATION NUMBER: US/08/852,824C

; CURRENT FILING DATE: 1997-05-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 384  
; TYPE: PRT  
; ORGANISM: genomic  
US-08-853-824-4

Query Match 55.6%; Score 5; DB 3; Length 384;  
Best Local Similarity 100.0%; Pred. No. 2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
| | | | |  
DB 6 TPVAP 10

## RESULT 44

US-08-769-309A-15  
; Sequence 15, Application US/08769309A  
; Patent No. 5741890  
; GENERAL INFORMATION:  
; APPLICANT: Scott, John D.,  
; APPLICANT: Nauekt, Brian J.,  
; APPLICANT: Knauck, Theresa M.  
; TITLE OF INVENTION: Protein Binding Domains of Gravin  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower/233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/769,309A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5741890and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/33451  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 396 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-769-309A-15

Query Match 55.6%; Score 5; DB 1; Length 396;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
| | | | |  
DB 44 STPVA 48

RESULT 45  
US-08-994-570-15

; Sequence 15, Application US/08994570  
; Patent No. 6090929  
; GENERAL INFORMATION:  
; APPLICANT: Scott, John D.,  
; APPLICANT: Nauekt, Brian J.,  
; APPLICANT: Knauck, Theresa M.  
; TITLE OF INVENTION: Protein Binding Domains of Gravin  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower/233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/994,570  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6090929and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/33451  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 396 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-994-570-15

Query Match 55.6%; Score 5; DB 3; Length 396;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 STPVA 6  
| | | | |  
DB 44 STPVA 48

RESULT 46  
US-08-839-008-5  
; Sequence 5, Application US/08839008  
; Patent No. 5916758  
; GENERAL INFORMATION:  
; APPLICANT: Hurle, Mark R  
; APPLICANT: McDonnell, Peter C  
; APPLICANT: McNulty, Dean E  
; APPLICANT: Rosen, Craig A  
; APPLICANT: Siemens, Ivo R  
; APPLICANT: Young, Peter R  
; APPLICANT: Yue, Tian-Li  
; TITLE OF INVENTION: Smooth Muscle Cell-Derived Migration Factor  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,008  
FILING DATE: 23-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/563,697  
FILING DATE: 28-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Baumeister, Kirk  
REGISTRATION NUMBER: 33,833  
REFERENCE/DOCKET NUMBER: P50384  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-270-5096  
TELEFAX: 610-270-5090  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 401 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-839-008-5

Query Match 55.6%; Score 5; DB 2; Length 401;  
Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 TPVAP 7  
Db 330 TPVAP 334

RESULT 47  
US-08-403-545-4  
Sequence 4, Application US/08403545  
Patent No. 5656483  
GENERAL INFORMATION:  
APPLICANT: Sokatch, John R.  
APPLICANT: Sykes, Pamela Joy  
APPLICANT: Madhusudhan, K.T.  
TITLE OF INVENTION: Genes Encoding Operon and Promoter for  
TITLE OF INVENTION: Branched Chain Keto Acid Dehydrogenase of Pseudomonas putida  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carolyn D. Moon  
ADDRESSEE: Dunlap, Coddling, Peterson and Lee  
STREET: 9400 N. Broadway, Suite 420  
CITY: Oklahoma City  
STATE: Oklahoma  
COUNTRY: USA  
ZIP: 73114  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette 5.25 inch, 360 Kb Storage  
COMPUTER: IBM AT  
OPERATING SYSTEM: MS-DOS Version 3.3  
SOFTWARE: Professional Write 2.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/403,545  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 7/603/781  
APPLICATION NUMBER: 07/172,148  
FILING DATE: 23-003-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Carolyn D. Moon  
REGISTRATION NUMBER: 33,022  
REFERENCE/DOCKET NUMBER: 5820.101

TELECOMMUNICATION INFORMATION:  
TELEPHONE: Attorney, (405) 478-5344  
TELEFAX: Attorney, (405) 478-5349  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 422 Amino acids  
TYPE: Amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: E2 component  
LOCATION: 3098-4363, does not include initiating methionine  
IDENTIFICATION METHOD: N-terminal sequence  
PUBLICATION INFORMATION:  
AUTHORS: Burns, Gayle  
AUTHORS: Brown, Tracy  
AUTHORS: Hatter, Kenneth  
AUTHORS: Sokatch, John R.  
TITLE: Comparison of the amion acid sequences of the  
TITLE: transacylase components of branched-chain oxoacid  
TITLE: dehydrogenase of Pseudomonas putida, and the pyruvate and  
TITLE: 2-oxoglutarate dehydrogenases of Escherichia coli  
JOURNAL: European Journal of Biochemistry  
VOLUME: 176  
PAGES: 165-169  
DATE: 1988  
US-08-403-545-4

Query Match 55.6%; Score 5; DB 1; Length 422;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VSTPV 5  
Db 351 VSTPV 355

RESULT 48  
US-08-404-381-4  
Sequence 4, Application US/08404381  
Patent No. 6168945  
GENERAL INFORMATION:  
APPLICANT: Sokatch, John R.  
APPLICANT: Sykes, Pamela Joy  
APPLICANT: Madhusudhan, K.T.  
TITLE OF INVENTION: Genes Encoding Operon and Promoter for  
TITLE OF INVENTION: Branched Chain Keto Acid Dehydrogenase of Pseudomonas putid  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Carolyn D. Moon  
ADDRESSEE: Dunlap, Coddling, Peterson and Lee  
STREET: 9400 N. Broadway, Suite 420  
CITY: Oklahoma City  
STATE: Oklahoma  
COUNTRY: USA  
ZIP: 73114  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette 5.25 inch, 360 kb Storage  
COMPUTER: IBM AT  
OPERATING SYSTEM: MS-DOS Version 3.3  
SOFTWARE: Professional Write 2.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/404,381  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/603,781  
APPLICATION NUMBER: 07/172,148  
FILING DATE: 23-003-1988



ATTORNEY/AGENT INFORMATION:  
NAME: Carolyn D. Moon  
REGISTRATION NUMBER: 33,022  
REFERENCE/DOCKET NUMBER: 5820.101  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: Attorney, (405) 478-5344  
TELEFAX: Attorney, (405) 478-5349  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 422 amino acids  
TYPE: Amino acid  
TOPOLOGY: Linear  
MOLECULE TYPE: Protein  
FEATURE:  
NAME/KEY: E2 component  
LOCATION: 3098-4363, does not include initiating methionine  
IDENTIFICATION METHOD: N-terminal sequence  
PUBLICATION INFORMATION:  
AUTHORS: Burns, Gayle  
AUTHORS: Brown, Tracy  
AUTHORS: Walter, Kenneth  
AUTHORS: Sokatch, John R.  
TITLE: Comparison of the amion acid sequences of the  
TITLE: transacylase components of branched-chain oxoacid  
TITLE: dehydrogenase of Pseudomonas putida, and the pyruvate and  
TITLE: 2-oxoglutarate dehydrogenases of Escherichia coli  
JOURNAL: European Journal of Biochemistry  
VOLUME: 176  
PAGES: 165-169  
DATE: 1988  
US-08-404-381-4

Query Match 55.6%; Score 5; DB 4; Length 422;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VSTPV 5  
|||||  
DB 351 VSTPV 355

RESULT 49  
US-08-353-550-1  
Sequence 1, Application US/08353550  
Patent No. 5744313  
GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Cavanaugh, William M.  
TITLE OF INVENTION: No. 5744313el Protein Domain Which Binds  
TITLE OF INVENTION: Tyrosine Phosphorylated Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourile and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/353,550  
FILING DATE: 09-DEC-1994  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 02307K-057300

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 428 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-353-550-1

Query Match 55.6%; Score 5; DB 1; Length 428;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 PVAPT 8  
|||||  
DB 361 PVAPT 365

RESULT 50  
US-08-551-687-1  
Sequence 1, Application US/08551687  
Patent No. 5925547  
GENERAL INFORMATION:  
APPLICANT: Williams, Lewis T.  
APPLICANT: Cavanaugh, William M.  
TITLE OF INVENTION: No. 5925547el Protein Domain Which Binds  
TITLE OF INVENTION: Tyrosine Phosphorylated Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Steuart Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/551,687  
FILING DATE: 01-NOV-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/353,550  
FILING DATE: 09-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy, Matthew B.  
REGISTRATION NUMBER: P39,787  
REFERENCE/DOCKET NUMBER: 2307K-5731  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 428 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-551-687-1

Query Match 55.6%; Score 5; DB 2; Length 428;  
Best Local Similarity 100.0%; Pred. No. 2.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 PVAPT 8

Wed Aug 15 13:34:53 2001

us-09-372-036-26.oli.ra1

Page 28

Db 361 PVAPT 365

Search completed: August 15, 2001, 12:33:43  
Job time: 48 sec

---